DOCUMENT RESUME

ED 101 322 CS 001 620

AUTHOR	Page, William D., Ed.
TITLE	Help for the Reading Teacher: New Directions in Research.
INSTITUTION	ERIC Clearinghouse on Reading and Communication Skills, Urbana, Ill.; National Conference on Research in English.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.
PUB DATE	Mar 75
CONTRACT	NEC-0-72-4636
NOTE	112p.
AVAILABLE FROM	National Council of Teachers of English, 1111 Kenyon Road, Urbana, Illinois 61801 (Stock No. 13915; $3.95 nonmember, $3.75 member)
EDRS PRICE	MF-$0.76 HC-$5.70 PLUS POSTAGE
DESCRIPTORS	Cloze Procedure; Content Reading; Elementary Secondary Education; *Informal Reading Inventory; *Miscue Analysis; Readability; Reading; *Reading Diagnosis; *Reading Research; *Reading Tests

ABSTRACT

This book is designed to be a bridge between teachers and research in such topics as the cloze procedure, readability formulas, miscue analysis, reading strategies, informal reading inventories, and concepts and reading in the content areas. "Children's Behavior While Reading" by H. Robinson serves as a bridge between Huey's early beginning and contemporary theory and research. "Oral Reading Analysis: A View of the Reading Process" by C. Burke and "Reading Strategy Lessons; Expanding Reading Effectiveness" by Y. Goodman both discuss the reading process as seen through miscue analysis. "The Relationship of Beginning Reading Instruction and Miscue Patterns" by J. DeLawter looks at patterns of miscues in children who have received different kinds of reading instruction. "A Different Look at Reading in the Content Areas" by K. Carlson looks at the differences in miscues produced in reading different content materials. "Literacy in the Classroom" by J. Bormuth discusses readability of texts and the pupils' ability to comprehend textual materials. "Auditory Discrimination: Differences versus Deficits" by P. Williams reports on research that isolates linguistic elements. And "Use of Informal Reading Inventories" by W. Page and R. Barr discusses the informal reading inventory. (WR)

Help for the Reading Teacher: New Directions in Research

William D. Page, Editor

National Conference on Research in English

ERIC Clearinghouse on Reading and Communication Skills
National Institute of Education

National Conference on Research in English Bulletin Series
Sara W. Lundsteen, editor

Library of Congress Catalog Card Number 75-1750
Published March 1975

National Conference on Research in English

ERIC Clearinghouse on Reading and Communication Skills
1111 Kenyon Road, Urbana, Illinois 61801

Printed in the United States of America

Editor, Susan B. Drake
Designer, Bob Bingenheimer
Cover Illustration, Patricia Anderson

Contents

Foreword

When a child's score from a standardized test is studied by the teacher, it often evokes such comments as, "Charles does better than that in reading," or "I could have predicted that score for Sue," or "That score does not indicate the type of work Tom does." These reactions are typical because the good teacher is a keen observer. In fact, the trend in education recently is to place more weight on "teacher observation" than on formalized testing. This book is directed to the practitioners in the classroom, to aid them in the process of observing students reading and translating these observations into an evaluation of reading performance. The teacher in this role is a diagnostician with an intellectual approach for identifying reading strengths and weaknesses in order to aid the student

A conference titled "What Kids Do in Reading" at the University of Chicago in summer 1972 provided a forum for airing new ideas and insights into observing the reading process. The papers collected here emerged from the research that conference generated and present ideas that can be applied directly in the classroom. Sponsored by the National Conference on Research in English (NCRE) and the ERIC Clearinghouse on Reading and Communication Skills, this monograph is brought to teachers under the editorship of William D. Page.

The ERIC system makes available—through the ERIC Document Reproduction Service—a wealth of information, including all federally funded research reports since 1956. If the findings of specific educational research are to be rendered intelligible to teachers and applicable to teaching, considerable bodies of data must be reevaluated, focused, translated, and molded into an essentially different context. Rather than resting at the point of making research reports readily accessible, the ERIC Clearinghouses commission recognized authorities to produce state-of-the-

art papers, bibliographies, and monographs in specific areas. The present publication is designed to be a bridge between teachers and research in such topics as the cloze procedure, readability formulas, miscue analysis, reading strategies, informal reading inventories, and concepts and reading in the content areas.

As the body of information derived from educational research has expanded, so has the gap between research and classroom teaching. But if teachers can apply the concepts presented in this monograph, the widening gap between research and the classroom will narrow and students will have an improved environment for developing their ability to read.

Sister Rosemary Winkeljohann
ERIC Clearinghouse on Reading and Communication Skills

Introduction

If Edmund Huey could read this book, he would find in it extensions of the theories and methodologies he began to explore more than three-quarters of a century ago. He might find it strange that reading research and theory wandered for so many years in a largely nonproductive wilderness before they came back to some essential realities:

Understanding reading must be built on understanding what happens when the reader is interacting with written language.

Insights based on responses to *bits and pieces* of language (letters, sounds, words, sentence fragments, lists of unrelated sentences) do not illuminate how readers get meaning from written language or learn to do so.

Reading is a receptive language process. It must be studied using tools and insights appropriate to linguistic research.

The reading process may be the same for everyone but each pupil learns to control it individually. Productive instructional decisions are best made by teachers who understand the process and know how to interpret pupils' use of that process at any given time.

Physical and mental abnormalities in learners may influence the acquisition of reading, but even abnormalities must be considered in a framework of how meaning is constructed from written language.

Full appreciation of reading as the receptive phase of communication between the writer and reader means awareness of the extent to which the specific conceptual background and interest of the reader influence comprehension.

Helen Robinson, whose paper serves as an historical bridge from Huey's early beginning to contemporary theory and research, quotes Huey's simple, direct challenge: "We have surely come to the place where we need to know just what the child normally does when he reads, in order to plan a natural and economic method of learning to read." He said that in 1910.

Commitments to knowing "what the child normally does" as a means of planning "a natural and economic method of learning to read" are the threads that run through the papers in this volume. Robinson comments that research has focused more on the material to be read than on the reader. The authors in this volume, however, look at both the reader and the written language in the context of the interaction between the two.

The papers of Carolyn Burke and Yetta Goodman plunge us into the reading process as it may be seen through miscues, the unexpected responses readers produce. Burke asks, and answers, the question, "How can we [teachers] gain the information needed to offer reading instruction without disrupting the reading process?" Her answer is to learn to use the miscues children make within a framework of how reading works. She warns us not to be superficial or too quick to make snap judgments, "to assume that all of the words the reader has omitted are unknown" or that certain substitutions indicate phonics problems.

Burke illustrates the reading process with examples of actual reading miscues and at the same time demonstrates how to observe and interpret miscues. For the teacher, Burke argues, miscue analysis is a way of "tapping the reading process as the reader uses it," so that the teacher can "support the reader in discovering and developing those techniques and understandings" necessary to independent flexible reading.

Goodman labels these "reading strategies." They are, she says, "natural to the reading process." "Reading strategies are the myriad ways readers process information when dealing with written language." The teacher's role, as Goodman describes it, is one of facilitating the development of these strategies. This is a sharp contrast to the traditional "skill" focus, in which teachers are viewed as the source of specific skills which they instill in their pupils.

Picking up where Burke leaves off, Goodman describes a young reader using "the semantic and syntactic system as well as a graphophonic system" in dealing with "natural written language." The teacher "must discover which strategies the reader is using effectively and which strategies the reader needs help to develop."

Reading strategy lessons based on evidence from miscue analysis are the means Goodman advocates for this supportive help, for this material is natural, meaningful, and designed to focus the readers' use of appropriate strategies. Goodman emphasizes the crucial "need for teachers knowledgeable about language, how language is learned, and what the reading process is like."

DeLawter explores patterns of miscues in children who have received different kinds of reading instruction. Her study shows that instructional focus in reading does affect the reading process and the types of miscues children produce, but that universal aspects of the process appear regardless of instruction.

Carlson looks at the differences in miscues produced in reading different content materials. He finds that his research and similar research show the reading process to be essentially the same across content materials. The focus, he feels, needs to be on building concepts appropriate to each area. Carlson suggests that miscue analysis be used to indicate the conceptual problems readers are having.

Bormuth draws on his work using the cloze test to examine the readability of texts on the one hand and pupils' ability to comprehend texts on the other. Readability formulas have dealt with quantifiable aspects of written language in order to assign a level of difficulty to any given text, but they have left out the learner. Comprehension tests, on the other hand, have been designed to assign a level of competence in reading comprehension to a reader which would then apply to anything he or she reads.

Bormuth distills from his research and that of others what he hopes will help teachers find appropriate reading materials for specific pupils which they can comprehend and from which they can learn. In this latter sense, Bormuth moves beyond common definitions of comprehension to one that involves acquiring new knowledge through reading.

Peggy Williams reports a study that exposes the misconceptions that can result from research which isolates linguistic elements. She finds that "evidence" linking "poor auditory discrimination" (as demonstrated on the Wepman test) with poor reading among black children is in fact largely a reflection of the inappropriateness of the test for the group. The performance reflects language difference, not auditory deficiency. Reading programs "resulting from inaccurate descriptions have consumed considerable time and money but failed to extensively improve the lot of the 'underprivileged' child," she warns.

In the concluding paper, Page and Barr bring together the

strands of the renewed interest in planning instruction by observing the reader interacting with written language. They build around the informal reading inventory, a simple device that any teacher or reading specialist can use. The authors show how new insights, new devices, and new understandings of the reading process can make the informal reading inventory a tool that will provide an insightful practitioner the information needed to plan effective personal instruction.

This book has many practical suggestions for instruction. All of them come from our return "to the place where we need to know just what the child normally does when he reads."

Kenneth S. Goodman
Wayne State University

1 Insights from Research
Children's Behavior while Reading

Helen M. Robinson

Scattered reports of children's behavior while reading were made before the turn of this century. Besides, every good teacher has records, either on paper or in memory, of the reading characteristics of some pupils. In 1910, Huey said "We have surely come to the place where we need to know just what the child normally does when he reads, in order to plan a natural and economic method of learning to read" (p. 9). With such a long history of observations, why are we returning to this topic as a modern issue?

At least five reasons appear to the writer to be of major importance. First, the best minds in the field of reading and in allied fields are challenged to throw light on the reading process. If the process could be understood, it would be possible to identify factors that facilitate as well as interfere with reading, from the beginning through the mature stages.

The best that we can do at present is to infer the process from the product. As a result, widely different interpretations can be made of the same sets of test scores, errors in oral reading, observations of children while they read, eye movement records, and other data. The point of view of the person who interprets these data often makes a marked difference in the conclusions reached. From such differences have come a host of models of reading or of the reading process. As a part of the targeted research sponsored by the United States Office of Education, Davis (1971) prepared and edited a lengthy report. The report shows that some models of reading represent a proliferation of skills, while other models are so simple that they embrace only a minute segment of what is usually considered reading.

A second reason for examining children's behavior while reading is that in a new era, children may be different than they were in earlier years. Travel and TV, for example, have created experiences and attitudes different from those of the rural child who was only

familiar with animals and family life. Furthermore, our attention has been focused on a group of children once overlooked—inner city children. The universality or diversity of reading behaviors may make more sense if a relatively good cross-section of the population is studied.

A third reason to reexamine children's behavior is that we have new foci for research. Teachers or researchers see essentially what they are looking for. What Gray (1916) called oral reading errors, Goodman (1968) calls miscues. The difference is not one of terminology alone. Whereas Gray looked for mechanical problems such as faulty vowels and consonants, reversals, addition and omission of sounds, and the like, Goodman focuses on successful or unsuccessful use of the cues. In particular, he is concerned with how the miscue approximates the text in terms of syntactic (arrangement of words in grammatical units), semantic (meaning), and symbol-sound variations. Taylor (1937) trained children and adults to eliminate regressive movements of the eyes mechanically; recent investigators explain that even the good reader must regress to examine the text if meaning is interrupted. Early studies found word-by-word reading to be characteristic of immature readers and urged that children be taught fluency from the beginning. Clay (1966) found that pointing to words is a necessary stage in learning to match printed to spoken words. The foregoing examples make it clear that recent orientations give new information because of what we see and hear as we study children.

A fourth reason to reexamine children's reading behavior is to quantify the data obtained from such observations. This is not to say that we must return to the complicated statistical procedures that often yielded little because the input was limited. Instead, case study data can be quantified to reveal characteristic behaviors at various stages of reading growth. Then atypical behavior patterns can be examined in relation to stages of growth and individuals with acute needs can be identified.

A fifth reason for studying children's behavior is to generate hypotheses for other types of investigations. An example is Clay's (1966) discovery that many beginners do not know the boundaries of printed words. Following her report came experimental research by a half-dozen persons who confirmed her findings and extended them to spoken words, letters, and sounds.

Historical Perspective

An historical examination of reading behavior would be incomplete without mention of Huey's famous report (1910) of the early

research and his interpretation of it. His data were based largely on retrospective reports of sophisticated readers and crude eye-movement measures. For teaching, the consequences of this early report were many. A strong impetus was given to the teaching of words prior to letters by conclusions such as "reading was in larger wholes than letters" (p. 73). Huey forecast the instructional program developed later and used predominantly until recent years: word form is composed primarily of the consonants since they protrude above and below the lines (p. 81); meaning dominates the perception of words and phrases (p. 116); stumbling and hesitation in oral reading comes from too much attention to the mechanics of reading, especially phonics (p. 302); "We know that the reading of life is almost exclusively silent reading. Yet in preparing for life, we are instructed almost exclusively in reading aloud . . ." (p. 10).

Furthermore, Huey wrote at length about the "natural" way of learning to read at home. Huey felt that, just as children's curiosity is satisfied by answering their questions, children would accumulate and learn better a large stock of sight words if their curiosity about print were satisfied. In an eminently modern suggestion, he stated that, "The secret of it all lies in parents' reading aloud to and with the child."

In summary, many of the materials and methods for teaching reading today were suggested in these early years and, as Huey said, the reason we continue to do many things is because that is what has always been done (p. 11). In this monograph, we are looking for better reasons for what is being done.

To study the reading process, two essential ingredients must be examined: the reader and the selection read. If the significant characteristics of each could be identified, then the interaction of the reader and the material could be interpreted. Far more research has been done on the materials than on the reader, although many inferences have been made from the materials and attributed to the reader. To determine the accuracy of the inferences, a great deal more must be learned about the reader.

Materials have been studied extensively because they are static and tangible. They lend themselves to word counts, syntactic analyses, and myriad alterations. In contrast, the readers change constantly, have been influenced by several years of home and neighborhood conditions, possess varying amounts and types of language facility, and differ in their interests in and motivations for reading. For this reason it has not been possible to isolate all of the individual factors bearing on reading behavior nor to assess the influence of combinations of factors on the interaction of the

reader with the materials. These goals must be achieved in order to obtain accurate knowledge of the reading process.

Approaches to Describing Reading Behavior

Introspection and retrospection were used widely in reading studies before the turn of the century. Thereafter, most of the procedures now being used were partially described or forecast, especially by Huey (1910). Some techniques attracted more attention and were more fully developed at an earlier date than others.

Testing procedures. Parallelling the study of eye-movements was the development and extensive use of tests. Reading tests gave vivid evidence of the wide ranges of reading achievement in most classrooms. From the first standardized tests before 1920, children's behavior in silent reading has been assessed by the answers they gave to questions about what was read. In the early years, teachers scored the tests and theoretically had the opportunity to see what kinds of errors were made. Increasingly, tests have been scored mechanically, with answer sheets and finally by machine. As test scoring became more mechanical, information about children's behavior decreased until nothing was known except an overall score or scores on parts of the test. This right-wrong scoring of answers became prevalent in classrooms as a means for teaching comprehension in silent reading. Few teachers took the time to ask children how they arrived at answers, much less to lead them through the steps in arriving at correct answers. Thus the teaching of comprehension became, in essence, an exercise in testing. Children who were unable to learn, largely by trial-and-error, to answer questions correctly, continued to practice their errors until they attained a proficiency in making errors that led to failure.

The Gray Standardized Oral Reading Paragraphs (1916) provided ample opportunity for examiners to hear and observe children's behavior while they read. However, the recording system forced attention to particular errors which were analyzed mechanically. Errors were largely explained as inaccurate perception and the remedy was drill, often with the words in isolation.

The testing movement had a number of consequences for teachers. First, silent reading tests diverted teachers' attention away from pupils' behavior toward overall scores. Since there was little helpful information in the scores, analyses of difficulties and strengths were left largely to teachers' observations. Moreover, few teachers were trained to make optimal use of their observations in order to help children solve their learning problems.

Second, the testing of large numbers of pupils revealed the wide ranges of reading achievement at any single grade level. Emphasis was placed on getting those below the norm up to it. The challenge of reading disability and retardation increased through the decades.

Third, most of the research in reading was based on performance on standardized reading tests. As scores of large numbers of pupils were analyzed statistically, the individual was lost in the group. Furthermore, nonstatistical research was considered less than scientific and, as a result, insights acquired from close observation of behavior were lost.

About 20 years ago, the cloze procedure, a type of completion test, was devised to measure readability (W. Taylor 1953). A number of investigators have shown that the cloze is an adequate test of comprehension (Bormuth 1969a). However, scores on the cloze have been validated against conventional tests and experimenter-made test questions of which Bormuth (1969b) has been exceedingly critical.

The cloze procedure has also been used for instruction in comprehension with persons from first grade to college. The results of using the test alone have not been promising except in two studies where instruction accompanied filling in the cloze blanks. Jongsma (1971) has suggested that instruction aimed at improving vocabulary or the use of context clues, if conducted over a sufficiently long time, might be useful at particular grade levels.

Like standardized reading tests, the cloze tests manipulate text but give little insight into the reader unless other techniques are added. However, they have the advantage of allowing the study of various linguistic classes of words, and as used by Jenkinson (1957) with introspective and retrospective techniques, permit many inferences about printed language.

Eye-movement photography. The invention of motion picture film before 1920 permitted extensive studies of the behavior of a person's eyes while reading varieties of texts. A host of research followed, comparing oral with silent reading, using different types of materials read for different purposes, and comparing eye-voice span. Buswell (1922) and his associates, for example, demonstrated a characteristic pattern of increase in span and length of fixations accompanied by a decrease in regressive movements from first grade to adulthood, with the most marked change occurring during the first four grades. Large individual differences were found to be related to performance on standardized tests.

Some very significant implications for teaching reading came from the eye-movement studies. First, there came an emphasis on

developing competence in silent reading rather than just in oral reading. Today silent reading is usually taught as early as the first year in school.

A second consequence was the impetus to use a wide variety of materials to teach reading. Since eye movements differed when children read in different contexts and subject areas, it was reasoned that children should learn how to read in science, social studies, and all curriculum areas.

A third consequence has been an emphasis on rate of reading, both for adults and children. At one time, some persons were convinced that reading rate could be improved by improving the mechanics of eye movements; that is, increasing the span with short exposures and decreasing the regressive movements by covering materials as they were read. Many investigators seemed to agree that motivation for rapid reading was the common element in all rate techniques. Furthermore, most studies supported Buswell's original conclusion that eye movement characteristics are symptoms rather than causes of slow reading rate.

A fourth consequence of the eye-movement studies has been a number of admonitions to teachers—i.e., discourage repetition, urge rapid reading, and eliminate pointing to words or you may reduce fixations and spans. Later in this paper we shall see that other interpretations suggest that these admonitions may be incorrect.

Today, eye-movement photography has been revived. Newer techniques such as the electromyograph permit much greater freedom of movement of the eyes. Furthermore, the computer can be used to report various characteristics of eye-movement behavior. These modern findings will be interpreted differently depending upon the orientation of those doing the research. Those concerned with information processing suggest that the mature reader uses fewer cues than the immature reader and, therefore, successive eye fixations may be determined by how closely the language of the text fits that anticipated by the reader. Additionally, even regressions have been found useful when they provide the reader with more cues to meaning (Goodman and Burke 1968).

Introspective-retrospective reports. In the last century introspective-retrospective reports were used by sophisticated subjects. However, the verbalization of unsophisticated subjects interpreted by psychologists is of relatively recent origin (Duncker 1945). Piaget's (1952) interpretation of introspective-retrospective reports by children is probably the best example. Swain (1953) applied this technique to college students' answering questions; Piekarz

(1954) searched for the most productive among thirteen ways of presenting materials. Subsequently, introspective-retrospective reports were used with subjects from third grade to college. The studies contrasted responses of good versus poor readers in a variety of situations: prose and poetry interpretations; reading for different purposes; and reading in different content areas. Highly creative students were compared with those having high I.Q.'s; students' involvement in reading was also explored.

The range of topics investigated was wider than could be explored by the techniques previously discussed. Furthermore, in most instances the learner was the focus of study rather than the materials.

However, three major difficulties are inherent in this technique. First, it is questionable whether subjects really verbalize all of their thoughts. Second, as Simons (1971) points out, mental processes that occur rapidly may not even be available to the subject for verbalization. Third, the mass of verbal materials must be analyzed so that the investigator obtains information from it. Therefore, schemes for classifying protocols may vary widely depending upon the orientation of the investigator.

Perhaps the most important implication of this type of study, for teachers, is the recurrent suggestion to ask students why or how they arrived at given interpretations. Students' reactions to what they read may be emotional or intellectual; that is, students may adopt strong positive or negative attitudes toward the materials and refuse to get the author's ideas, or they may misinterpret the author's ideas because they fail to follow the logic of an argument. Whatever the problem, teachers who ask students to verbalize reasons for their responses to selections find the experience rewarding.

Systematic observations and descriptions. In the last decade, researchers have returned to observing and describing children's behavior while reading. These recent reports emphasize systematic observations rather than random or casual ones. The number of subjects observed has increased as has the period of time over which observations are recorded.

Notable among the investigators have been MacKinnon (1960), Clay (1966), Goodman and Burke (1968), and Söderberg (1971). All except Clay's work have been published, so it will be summarized briefly. She selected 100 children entering five schools in New Zealand with a mean age of 5 years, 1 week. Once every week of the academic school year each child read, to the investigator, the pages of reading materials for which the classroom teacher had

prepared him or her. Clay used a check sheet derived from a pilot study the preceding year. In addition, she made appropriate notations, especially during the pre-reading period. Children's comments were noted. Generally, the "most strategic observations" were ordered to be sure that common data were collected. For example, first priority was given to all oral responses to written text, then to motor responses, rate of reading, and any indications of feelings.

In addition to observations, tests were given at mean ages of 5 years, 5 years 6 months, and 6 years. The earliest tests were of language, auditory memory, and perceptual correlates. The Metropolitan Reading Readiness Test was given at all three ages. The criterion test was a combination of a word perception test devised by Clay and the Schonell R_1. Finally, the book being read and the child's accuracy level on this book were considered. Based upon achievement at the end of the first year, the pupils were divided into quartiles so the accumulated data could be studied in relation to achievement.

All of the children had a period of reading readiness and then were given books with instruction stressing fluency, meaning, and "learning as one reads" with slight attention to letter-sound associations or a basic sight vocabulary. At first, children drew pictures, created captions, and read them. When they performed this task well they were promoted to reading readiness books, then on to preprimers.

Clay observed that behavior during the preparation period appeared to fall into three major categories. Toward the end of the preparation period, the three categories seemed to be combined into matching behavior. The first category, visual sensitivity to letter and word forms, was demonstrated largely by children's scribbling, which seemed to show awareness of what print represented. The developmental sequence for letters, words, and word groups was scribble, invention, realistic copy plus invented variations, errors within almost correct patterns, and, finally, correct forms. Among the 100 children studied, some could already print their names when they entered school but others required an entire year to do so.

Directional orientation was the second category: to observe this, Clay asked children to point with their fingers while they read. They used their left hands more often than their right hands, which Clay felt was due to the fact that in the books used print was on the left-hand pages and pictures were on the right. The children's early behavior showed left-to-right, right-to-left, top-to-bottom, and

bottom-to-top movements as well as snaking movements over the lines. One child proceeded right-to-left on left pages and left-to-right on right pages. In general, the upper half of the group acquired consistent directional movements in 7 to 12 weeks; the lower achievers required about half a year and three pupils had not yet acquired this behavior by age 6.

The third category, speech behavior, concerned what was said about picture captions. The steps observed were naming, a repetitive short sentence (Here is a ——), a statement appropriate to the picture but not the exact text, memorized text, and finally a word-by-word response which might be inaccurate.

Integration of the three foregoing categories into matching behavior occurred before children were judged by their teachers to be ready for their first book. Looking at print, children made speech responses, then checked whether the print matched. The following techniques were used: page matching, line matching, word and letter concepts, locating specific words when the text was known, reading the spaces or voice pointing (staccato response), correction due to mismatching the number of words, beginnings of word control (those the child had learned), and self-correction when visual-vocal mismatches occurred. Self-correction was manifested by 90 percent of the children within 3 weeks of their promotion to books.

Of special interest was the sequence of behaviors related to fluency. At first, children reading captions were as fluent as in their oral speech, with no awareness of errors. As they began matching behavior with fingers pointing to printed words that matched the text, each word was emphasized, in a staccato fashion. Transition from finger pointing to voice pointing led to lighter stress of the breaks between words and finally to phrasing.

At the book-reading stage, a number of significant findings were noted. Among them: the average pupil in the high group read about 20,000 words per year in contrast to 5,000 words for the average pupil in the low group; pupils in the high group averaged 1 error in 37 words contrasted with 1 in 3 for the low group; the high group corrected 1 in 3 errors while the low group corrected 1 in 20 errors.

Clay's Conventions of Written English Test, given at age 5.0, was the best predictor of reading achievement at age 6.0 ($r = .79$). This test includes 20 items and should be most useful to teachers, with its predictive value overshadowed by the diagnostic value.

The findings of this extensive observational study are important in themselves, not to mention that hypotheses could be generated

from the study for a lifetime of experimental research. A few of the implications of the study for teachers are worth mentioning. First, children need to learn the direction of English print. Too often teachers take this characteristic for granted. When one-fourth of an experimental group has not learned to follow print, left-to-right, to return to the beginning of the next line, and to move from one page to the next in order, after six months, it is a significant source of reading difficulty. It is equally obvious that teachers can stress this convention of English in many ways.

Second, it is important to determine whether children know what is meant when the teacher talks about a word or a letter. Since only 47 percent of the pupils in Clay's study could isolate "any word" (not pronounce it) and 48 percent "any letter" at the end of one year in school, it is obvious that discussions of letters and words confused at least half of these children. Several studies done since (Downing 1970; Meltzer and Herse 1969) have confirmed this finding for British and American children. Direct instruction should be given early and consistently until children attain the concept of a word and a letter.

Third, finger pointing to words and a staccato pronunciation may be an important stage in matching printed words to spoken words. Children in the early stages of learning to read may need this crutch. For years we have discouraged, or not even permitted, pointing to words, thereby depriving children of the prop they needed to move to the next stage. It seems likely that pupils will stop using their fingers and discontinue staccato pronunciation as they acquire competence in reading. If they do not, then teachers can discourage a habit no longer needed.

Closely related is the mistaken expectation of fluency in reading from the very beginning. Since fluency may result from memorization and diminish to "word calling" during the early matching stage, it should not be expected. If fluency does not reappear at a later stage, it can be encouraged without fear of halting pupils' progress.

A fourth implication applies to all of the early reading behavior of children. Clay says of the pupil: "He must be given time to respond and must not be harassed because he searches at length, or because he fails to respond" (1967, p. 26). Too often the rush of "covering ground" causes teachers to tell children the words, or to permit able readers to make most of the responses. This practice may cause less mature readers to give up their search for clues and become dependent on their teachers. Furthermore, these children lose the opportunity to practice new approaches which they can

use later. An easy pace should be accompanied by help and guidance as needed to avoid unsuccessful behavior or continuous practice of errors.

Finally, Clay suggests that a number of the so-called reading readiness attributes may be acquired along with learning to read. In some schools, the lower half of the class may begin to read late in the school year and because of their slower responses, get very little practice in reading. It seems ironic that the children who need the most practice in reading really get the least.

The foregoing study and other observational investigations record behavior which is subject to different interpretations. Clay showed how her data fit the theoretical frameworks of cognitive psychologists and psychologists who have written about perception as well as fit information processing models and linguistic models. The data on behavior partially fit different theories, and it is obvious that researchers with different orientations could easily arrive at divergent conclusions from the same data. Inferences about the reading process will thus differ accordingly.

Teaching retarded readers. The behavior of retarded readers has been observed and recorded in detail for many years. Reading clinics, such as at the University of Chicago, where daily diary records are kept, offer unlimited opportunity to examine reading behavior. The major problem in interpreting the behaviors is the assumption that something fundamental is wrong with the learner. In some instances, this assumption may be accurate. However, it is possible instead that learning was delayed at different stages of development by inappropriate instruction, prolonged school absence, or personal problems that distracted the child's attention from reading.

Numerous retarded readers have succeeded in learning to read, a success which points out the importance of motivation. The young person or adult who seeks help is usually highly motivated when he or she begins instruction. Others need motivation, both from the teacher and from having access to relevant and readable materials.

Major research efforts have focused on the learner and involved teams of specialists (Robinson 1946 and others afterward). In addition, diary records give attention to materials. The importance of such observations will be great once enough data are secured on pupils who make progress, for only such data will allow an understanding of the stage at which the learning of retarded readers is halted.

Summary

Reading personnel are engaged in an intensive search for information on the reading process. Each technique discussed has been assessed in terms of information elicited by the technique about the learner and the materials. If the techniques were placed on a continuum, standardized tests and the cloze procedure fall at the materials end. Eye-movement photography ranks close to tests but offers some insight into personal behavior. Remedial reading instruction places closer toward the end labelled "learner." Observations of behavior such as Clay's describe the materials but emphasize the learner. All techniques require interpretation of the data, interpretations that are largely dependent on the theoretical biases of the person making the interpretations.

Each research technique has strengths and weaknesses. The best approximation to understanding the reading process should come from combining techniques whose strengths compensate for each others' weaknesses. No doubt reading research is still in the early stages of scientific development. As Simons (1971) wisely says, facts are being gathered but the time has not yet come when relevant facts can be separated from the irrelevant.

Research might profitably use groups of children with different backgrounds, languages, and cultural motivations to learn to read. Planned by a research team representing various theoretical orientations, observations of pre-reading behavior and longitudinal observations of children as they learn to read, using different approaches and different types of teachers, should provide a data bank for fruitful interpretation. Those who have analyzed materials for reading might supply the "best" available. Teachers would need to be trained to use these materials with the cloze procedure, with Bormuth's "Wh-questions" (1969b), with the best conventional questions that could be posed. Answers should be accompanied by introspection and retrospection concerning ways of arriving at answers or filling in cloze blanks. Opportunities for both silent and oral reading should be provided, so that miscues from the latter could be determined. Exploring pupils' concepts of what reading is and does (Estes 1971) and examining the internal motivation to correct miscues (Cohn 1972) could be useful. Pupils who experience unusual difficulty could be studied individually and careful records made of their learning behavior. Eye-movement records of samples of pupils could complete the data. The data collected could be submitted to experts with different orientations. From the common core of inferences supplemented by unique interpreta-

tions might come a productive model around which future research could be organized.

Meanwhile, teachers in the schools will continue to use the most relevant insights into children and the best materials they can find to produce literate youth who can read and who want to use reading.

REFERENCES

Bormuth, John R. "Factor Validity of Cloze Tests as Measures of Reading Comprehension Ability." *Reading Research Quarterly* 5 (Spring 1969a): 358–67.

———. "An Operational Definition of Comprehension." In *Psycholinguistics and the Teaching of Reading*, edited by Kenneth S. Goodman and James T. Fleming. Newark, Delaware: International Reading Association, 1969b, 48–60.

Buswell, Guy Thomas. *Fundamental Reading Habits: A Study of Their Development*. Supplementary Educational Monographs, No. 21. Chicago: University of Chicago Press, 1922.

Clay, Marie M. "Emergent Reading Behaviour." Ph.D. dissertation, University of Auckland (New Zealand), 1966.

———. "The Reading Behaviour of Five Year Old Children: A Research Report." *New Zealand Journal of Educational Studies* 2 (May 1967): 11-31.

Cohn, Marvin. "The Reader's Rudder." *Elementary English* 49 (March 1972): 369-71.

Davis, Frederick B., ed., "The Literature of Research in Reading with Emphasis on Models." Project No. 2: The Literature Search, Contract No. OEC-0-70-4790 (508), Project No. 0-9030. Graduate School of Education, Rutgers, The State University of New Brunswick, New Jersey, 1971.

Downing, John. "Children's Concepts of Language in Learning to Read." *Educational Research* 12 (1970): 106-12.

Duncker, Karl. "On Problem-Solving." Translated by L. S. Lees. *Psychological Monographs*, 58/5 (1945): iii-113.

Estes, Thomas H. "A Scale to Measure Attitudes toward Reading." *Journal of Reading* 15 (November 1971): 135-38.

Goodman, Kenneth S., and Carolyn L. Burke. "Study of Children's Behavior While Reading Orally." Final Report, Project No. S 425, U.S. Department of Health, Education and Welfare, 1968.

Gray, William S. *Standardized Oral Reading Paragraphs*. Bloomington, Illinois: Public School Publishing Co., 1916.

Huey, Edmund B. *The Psychology and Pedagogy of Reading*. New York: Macmillan Company, 1910.

Jenkinson, Marion Dixon. "Selected Processes and Difficulties of Reading Comprehension." Ph.D. dissertation, University of Chicago, 1957.

Jongsma, Eugene. *The Cloze Procedure as a Teaching Technique*. Newark, Delaware: International Reading Association, 1971.

Mackinnon, A.I. *How Do Children Learn to Read?* Toronto: Copp Clarke, 1960.

Meltzer, Nancy S. and R. Herse. "The Boundaries of Written Words as Seen by First Graders." *Journal of Reading Behavior* 1 (1969): 3-14.

Piaget, Jean. *Origins of Intelligence in Children*. Translated by Margaret Cook. New York: International Universities Press, 1952.

Piekarz, Josephine A. "Individual Differences in Interpretive Responses in Reading." Ph.D. dissertation, University of Chicago, 1954.

Robinson, Helen M. *Why Pupils Fail in Reading*. Chicago: University of Chicago Press, 1946.

Simons, Herbert D. "Reading Comprehension: The Need for a New Perspective." *Reading Research Quarterly* 6 (Spring 1971): 336-63.

Söderberg, Ragnhild. *Reading in Early Childhood*. Stockholm, Sweden: Almquist and Wiksell, 1971.

Swain, Emeliza. "Conscious Thought Processes Used in the Interpretation of Reading Materials." Ph.D. dissertation, University of Chicago, 1953.

Taylor, Earl A. *Controlled Reading*. Chicago: University of Chicago Press, 1937.

Taylor, Wilson L. "Cloze Procedure: A New Tool for Measuring Readability." *Journalism Quarterly* 30 (Fall 1953): 415-33.

2 The Reading Process

Oral Reading Analysis: A View of the Reading Process

Carolyn Burke

Reading as an Independent Experience

Reading is neither a team sport nor a spectator sport. The action occurs solely, and usually silently, between the author and the reader. In an attempt to interpret the author's message the reader makes use of his or her language, thought, and fund of knowledge to reflect on the language structure and thought of the author. As reading teachers, we can't sit figuratively on the sidelines to either call or analyse all of the "plays." Unlike football or baseball, each "play" or strategy applied by the reader is not strung out across a field for our immediate observation. There are no called "time outs" when "play" halts while the reader goes into a huddle and maps out strategy.

Because reading is, in many ways, a very private, personal, and independent experience it presents us, as teachers, with a very unique problem. How can we supervise reading without interfering with it? How can we gain the information needed to offer reading instruction without disrupting the reading process?

The Teacher as Participator

Our past attempts at offering reading instruction have focused on our participation as an active third party. Unfortunately, an interruption, no matter how momentary, disrupts the reading process. If we "give" the reader a word, we have failed to give him or her any way of handling, unaided, the next unknown item. If we stop the process and lead the reader through word attack strategies, we have so disrupted the reading that he or she is unable to employ the most useful of those strategies, context cueing. If we ask the student to re-read a section of text to hunt out a concept which he or she has failed to grasp or has interpreted in an unexpected manner, we tend to support the notion that there is one

right answer which must be and can be derived from the process.

Our very presence, as the third party, tends to make the neophyte reader psychologically dependent upon outside support. It leads both the student and the teacher to conclude that the student does not have the resources to handle reading problems and that any unsolved problems will lead to the total breakdown of the reading process.

As reading instructors, what we have found most perplexing is that the calculations and observations made by the reader are silent and unobservable. Even when we eavesdrop by having the student read orally, we are frustrated because we are observing only the performance of the whole process and not the process itself.

Oral reading is fleeting. The sounds flow continuously by our ear and are almost immediately beyond our recall. As a third party participator we are constantly forced to make snap judgements on the basis of segmented and limited surface level behavior. We find ourselves organizing and grouping phenomena on the basis of least resistance. We group and treat reading variations on gross and highly visible surface similarities. We begin to assume that all of the words the reader has omitted are unknown to him or her; that reading substitutions which have minimal sound or graphic similarity to the text item indicate that the reader does not effectively use phonic cues; that the presence of uncorrected items indicates a loss of meaning; that the reader should be consciously aware, as we are, of all his or her variations from the text.

Tapping the Reading Process

Instead of attempting either to participate in or to monitor the reading of our students we need to find a way of tapping their reading. We need a procedure which will allow our students to read uninterrupted and totally independently so that we may evaluate their use of strategies, their flexibility, and their effectiveness as readers. We need a procedure which will allow us to retrieve and examine the process, not the performance, and which will allow us the leisure for an in-depth examination of that process.

We need, in reading, a procedure which will match in usefulness the system of extended notation in mathematics. When our concern in a math class is not whether the student has arrived at the expected answer but whether he or she has proceeded through the necessary logical relationships, we ask the student to write out all of the calculations in solving a problem. The student can go through the process unaided and uninterrupted. At any later time

we can examine the calculations, determine what mathematical processes were used, note how the student related them to each other, and decide whether further instruction might be needed.

We need to find a way not to be team members or spectators but to be a part of the coaching staff, knowing that coaches are restricted from the playing field.

In reading there are no extended notations which remain as cues to the processing that has occurred. There are, however, oral reading variations which can be used for this purpose.

All readers have instances in which they produce unexpected responses, instances in which they vary from the printed text. These miscalculations are reached via the same reading process and the same strategies and language systems as the expected responses the reader produces. These miscalculations or miscues can serve the same function in reading that extended notation serves in math. Examining a miscue allows us to pinpoint the possible involvement of any and all of the related language systems. We can tally not only their occurrences but their interrelationships. Miscues which have a surface level similarity might or might not prove to have similar underlying causes. Miscues which are dissimilar on the surface can prove to have similar underlying causes. Analysis of reading miscues allows us to tap the reading process without interfering with it.

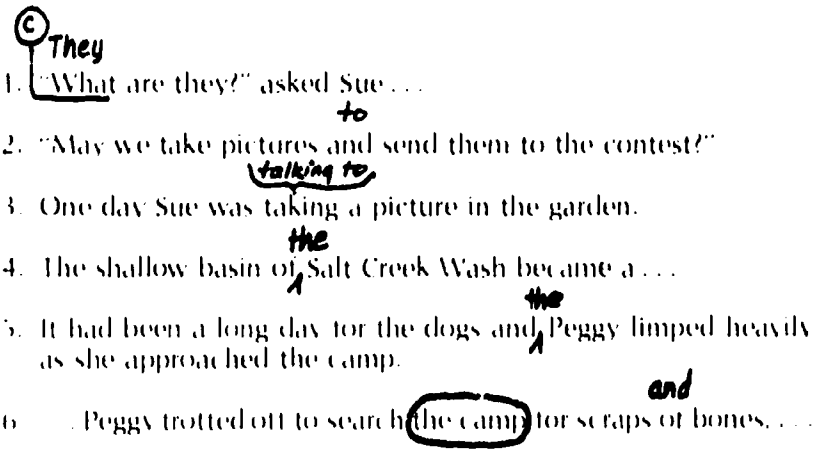
The process for collecting reading miscues is very simple. The reader needs to be provided with a text which is unfamiliar and relatively difficult for him or her to read. The reading should be tape recorded and uninterrupted and the students should retell the story onto the tape immediately following the reading. At least three perspectives are open to us through such miscue analysis: the reader's active role in the reading process, the effect of the material upon the reader, and a view of reading as a developmental process.

The Reader's Role

Readers have a very active and productive role in the reading process. The extent and nature of their activity is highlighted in their miscues.

In examples 1-6 the readers have produced a structural organization for the text sentence that differs from what the author produced. In some instances the reader was highly successful at changing the structure and was able to come up with a completely acceptable alternative: *Peggy trotted off to search for scraps and bones* and *May we take pictures to send them to the contest?*

In one instance a grammatically acceptable structure was produced which had a questionable or highly unusual meaning: *One day Sue was talking to a picture in the garden.* In still other instances the reader was unable to successfully complete the structure and was left with just the beginning portions of possible English sentences: *They . . .* and *It had been a long day for the dogs and the . . . .*



Just a quick look at these examples makes one aware that a number of factors might be involved in the readers' miscues: the occurrence of *they* in the near visual periphery in the first example; the close graphic and sound relationship between *taking* and *talking* in the third example; the compound noun structure of *the dogs* and *Peggy* in the fifth example. The complexity of factors serves to give each of the miscues a different surface flavor and helps to account for the varying degrees of success the readers had. But underlying all of these variations is the basic fact that the readers were actively anticipating and producing structures as they read. They were not passively taking in the author's prestructured material.

In much the same way as they anticipate structure, readers can be seen to anticipate meaning.

In example 7 the reader anticipates a common phrase, *things to do,* and then on the basis of the new coat and the graphic similarity between *horse* and *house* expects *rugs for our house* in place of *feed for our horse.* Living in an urban setting probably also contributed to this miscue.

The reader can be so sure of the story or plot line that he or she knows the intent of what is coming next and can produce such physically gross changes as *"Yes, you may go and have fun"*

(example 8) without changing meaning at all. Much the same situation exists in example 9, with the added element that the reader anticipated the repetition of structure from the prior sentence. Structure repetition is a likely possibility, as any connoisseur of primerese is well aware.

Example 10 comes from a story about a sick zoo elephant who refuses to take her medicine. One of the readers reacts to the fact that powders can generally be said to be *sprinkled* into something but only liquids can be *spilled*. The second reader, concentrating on the stealth of the doctor, *slipped* the powder into the water.

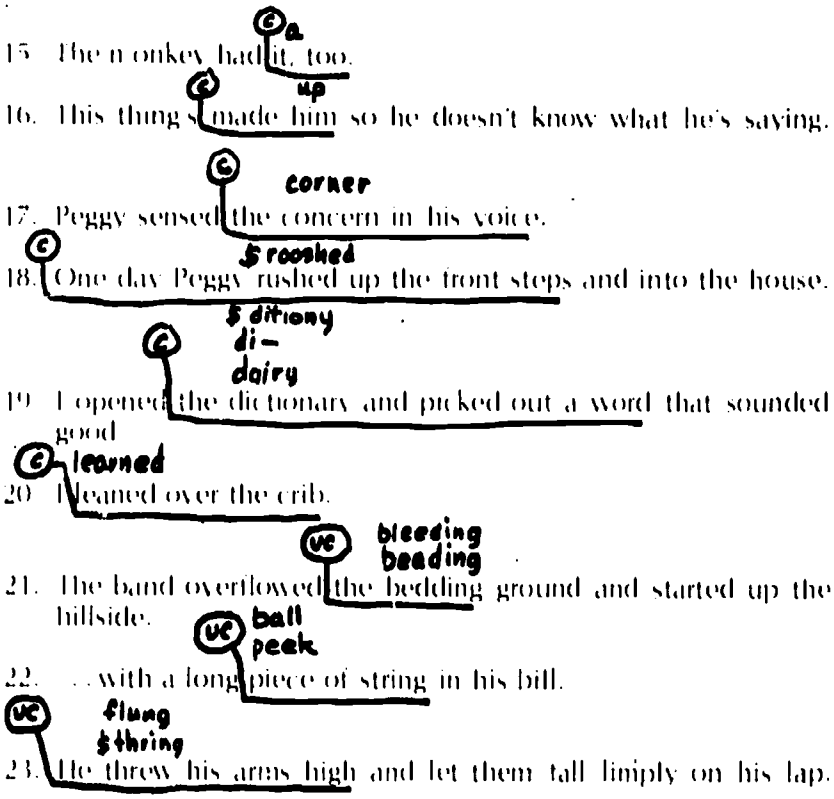
7. I wish we could all go.
   do
   But we have many things to buy.
   We have to get a new coat for you.
   rugs house
   We have to buy feed for our horse.
   may go and fur
8. "Yes, you do not have to stay home, . . ."
9. Away they went to the show.
   for the show
   They were not too late after all.
   S^2 slipped
   S^1 sprinkled
10. He spilled the powder into Sudana's water trough.
    learned
11. Cat laughed.
    aspirin
12. Let's try giving Claribel some oxygen.
    cyclamates
13. But I guess I added too many chemicals to the mixture.
    Taping
14. Tapping the reader.

One folk tale tells of the time when Tiger was not a good hunter and Cat agreed to teach him. Cat teaches and Tiger learns fast. Then when he thinks he has learned everything, Tiger decides to eat Cat but Cat has saved one trick. The Tiger says *"I do not know how to jump backward"* and the reader says Cat *learned* in place of Cat *laughed* (example 11).

Can we be surprised that in reading a story about a sick canary any modern child would suggest an *aspirin* (example 12)? Or, that in reading about a boy conducting experiments in his mother's kitchen, the reader might conclude that there were too many cyc*lamates* in the mixture (example 13)?

Nor was it surprising that a colleague from my office, who works constantly with the taped readings of children, should have read a heading in the rough draft of this paper and concluded that I mispelled *taping* (example 14).

These readers brought to the reading their knowledge of the world and anticipated correctly that it would help them interpret the messages the author had for them.



15. The monkey had it, too.
16. This thing s made him so he doesn't know what he's saying.
17. Peggy sensed the concern in his voice.
18. One day Peggy rushed up the front steps and into the house.
19. I opened the dictionary and picked out a word that sounded good
20. I leaned over the crib.
21. The band overflowed the bedding ground and started up the hillside.
22. . . with a long piece of string in his bill.
23. He threw his arms high and let them fall limply on his lap.

Whenever the reader anticipates what others do and recreates structures on the basis of these guesses, there have to be times when the reader is wrong, or when the alterations disrupt the structure. At these points the reader has the opportunity to pick up grammatical and meaning cues from the text and alter his or her guess.

In examples 15 and 16 the reader anticipated a different sentence structure: a direct object in the first instance and a verb + verb participle in the other. Corrections in such instances tend to be made almost immediately. The reader processes the next portion of text, realizes that it fails to fit structurally with what he or she has already produced, and corrects.

Several examples retain the grammatical structure of the author but involve semantic disparity. The length of the regressions or repetitions is closely related to the point at which the reader picks up the semantic cue necessary for correction and to the phrase structure within which the miscue occurs. *Voice* cues *concern, steps* cues *rushed, word* cues *dictionary*, etc. (examples 17–20).

Recognizing that you have miscalculated is usually, but not always, sufficient information for a successful correction. In example 21 the reader is aware that what he's producing doesn't seem to make sense but is unable to come up with the expected response. This is an average student in the upper grades. His problem is not one of word analysis on *bedding* but a conceptual one concerning his limited knowledge of sheep herding. In the next two examples (22 and 23) the reader picks up the necessary semantic cue but supplies an effective alternate.

Readers are active producers of language. They act upon the structures and meanings of the author on the basis of their own language and knowledge.

The Material

There is no real line that can be drawn between the influences and activities of the author and those of the reader within the reading process. But we can note that if both the reader and the author are active contributors to the process the possibility of miscues will increase at those points where the two differ. These differences can involve structural and word level variations between and within dialects. These variations need not cause a change in meaning.

24. . . . the headlamps (headlights) of the car . . .
25. Wait a moment. (minute)
26. He had a hypodermic syringe (needle) in one hand.
27. They put the hoses on the fire truck. (engine)
28. I was going to be ready to cut the bitten (bit) place and suck the venum out

Headlights for *headlamps*, *minute* for *moment*, and *made* for *had* might be considered lexical item variations across dialects, with *bit* for *bitten* involving structural dialect variations (examples 24, 25, 29, and 28).

made
29 The doctor had a guess.

down
30 Peggy gulped the biscuits and looked to the herder for more . . .

s
31 Something was wrong inside those four tons of flesh and bone.

I
32 I heard a musical whistle near my ear and thought it had come over the radio.

me
33. Please pass the clock

In the substitutions of *needle* for *syringe* and *engine* for *truck*, both options might be available in the reader's dialect, with the miscue representing the reader's preference (examples 26 and 27). At the structural level, *gulped down* for *gulped*, *bones* for *bone*, *I thought* for *and thought*, and *pass me* for *pass* are all surface level variations of the author's sentence (examples 30–33). In other words, both the author's sentence and the reader's sentence are transformed from the same deep structure.

The significant point here is that the author and the reader can agree upon the meaning and yet produce alternate structures. The mere occurrence of a miscue does not insure a meaning change or even a confusion over grammar.

The Developmental Process

Differences other than dialect or preference can occur between the author and the reader: use of unusual or complex structure by the author, limitations on the developing strategies of the reader. One interesting way to focus on these developmental differences is to look at the similarities and differences between miscues made by a number of readers on the same piece of text.

Examples 34 and 35 were both made by readers in the primary grades. They indicate how conditioned these readers have already become to the well-established patterns of their reading material, even to the point of expecting character names such as *Little Mitten*.

Examples 36 and 37 indicate the varying degrees of success with which a number of fourth graders dealt with *rough* and *trough*—words with a spelling pattern the readers found difficult. A closer look at *water trough* indicates the complexity of problems they faced. The phrase structure of noun adjunct + noun is not a frequent one and some of the readers anticipated a verb in the noun position. Those who anticipated a noun seemed to find the *ough* spelling pattern a problem. Capping this off is the fact that,

for these urban readers, *trough* is probably a totally unfamiliar item so they had few or no semantic cues available for correcting.

"I would"

34. No one said a word
said Little Mitten

35. I can help the little kitten.

s³ broogh
s² right
s¹ raw

36 . . . the length of her rough leg . . .

s⁵ truck
s⁴ full
s³ though
s² through
s¹ throw

37. I filled her water trough twice.

s² big
s¹ hug

38. Her huge ears were folded back against her neck and shoulders.

grains each

39. At sixty grams of sulfa for every thousand pounds of elephant,

grains

it would take 480 grams of sulfa.

Examples 38 and 39 indicate the alternatives available to a reader who meets an unfamiliar item. An over concern for graphic and sound similarity can result in the meaningless phrase *hug ears*. A focus on the larger context and meaning can result in the production of *big ears* as a useful synonym for *huge ears*. Similarly, the student reading *grams* as *grains* has indicated that while he doesn't recognize the term he has "gotten the message."

The final three miscues all show groups of readers faced with sentence structures which they find either unusual or complex. The structures *One day at breakfast* (example 40) and *After the cut in his allowance* (example 42) seem to be difficult due to the placement of the prepositional phrase. The readers are probably much more used to finding it following both the noun and verb phrases.

s³ a
s² the
s¹ he

40. One day at breakfast his father said, . . .

41. But he still thought it more fun to pretend to be a great scientist, mixing the strange and the unknown →

42 After the cut in his allowance, . . .

After the cut in his allowance, . . .

After the cut in his allowance, . . .

After the cut in his allowance, . . .

After the cut in his allowance, . . .

Similarly, these readers did not anticipate *strange* and *unknown* functioning in the noun position in the phrase *mixing the strange and the unknown* (example 41). They frequently ran this sentence into the next one or left it hanging with no final intonation in their search for the expected noun.

Using Miscue Analysis

These reader miscues have offered only the briefest sample of the possibilities that miscue analysis opens up for evaluative purposes. Furthermore, although the use of miscue analysis offers us a way of tapping the reading process and evaluating the efforts of the reader, it does not automatically give us a suitable teaching procedure.

We must begin to distinguish between reading and reading instruction. Reading, as an active language process between the reader and the author, must occur uninterrupted. During the process readers are on their own. They must handle problems independently, making the best use of the reading strategies they have available.

Reading instruction is intended to focus on the use and availability of strategies. The purpose is not to find out *Who went to the store?* or *Why was the baby crying?* or *What did Mrs. Thomas have in mind when she bought that basket of red peppers?* The purpose is to support readers in discovering and developing techniques and understandings that will extend their independence and flexibility in handling a variety of materials. By selecting appropriate text segments and producing specifically focused teacher-made materials, we can highlight the various aspects of the reading process for the reader's scrutiny.

To restate the point: Miscue analysis can make available to us the information upon which reading strategy lessons are based. Strategy lessons allow the reader to explore specific aspects of

the reading process. The success of our analysis and our instruction is found at the point where the reader independently participates in the reading process.

REFERENCES TO MATERIALS READ BY STUDENTS

1, 2, 3. "Kitten Jones." *Beyond Treasure Valley* (3.1), Betts Basic Readers. New York: American Book Company.

4, 5, 6, 17, 21, 30. "Sheep Dog." *Widening Views* (8), Sheldon Basic Reading Series. New York: Allyn & Bacon, Inc.

7, 8, 9. "A Lot to Tell." *Around Green Hills* (1), Betts Basic Readers. New York: American Book Company.

10, 31, 36, 37, 38, 39. "Zoo Doctor." *The Magic Word* (4), Macmillan Reading Program. New York: Macmillan Publishing Company, Inc.

11, 14, 18, 22, 34, 35. Experience stories or teacher-produced materials.

12, 32. "Space Pet." *The Magic Word* (4), Macmillan Reading Program. New York: Macmillan Publishing Company, Inc.

13, 33, 40, 41, 42. "Freddie Miller, Scientist." *Adventures Here and There*, Betts Basic Readers. New York: American Book Company.

15. "The Big Surprise." *Around Green Hills* (1), Betts Basic Readers. New York: American Book Company.

16, 24, 25, 26, 28, 29. "Poison." *Adventures in English Literature*. New York: Harcourt Brace Jovanovich, Inc.

19, 20, 23. "My Brother Is a Genius." *Adventures Now and Then*, Betts Basic Readers. New York: American Book Company.

27. "A Day At Home." *Reading Miscue Inventory: Readings for Taping*. New York: Macmillan Publishing Company, Inc.

Reading Strategy Lessons: Expanding Reading Effectiveness

Yetta M. Goodman

Reading strategies are the myriad ways readers process information when dealing with written language. They are the methods readers use to construct the message as they perceive the author produced it. Reading strategies are natural to the reading process and occur continuously as the reader strives to construct the author's meaning.

Some of the universal strategies which readers constantly apply include: (1) selecting graphophonic, syntactic, and semantic cues; (2) predicting graphophonic, syntactic, and semantic information; (3) confirming those predictions by asking themselves whether the results of the predicting strategies produced acceptable language; (4) applying correction strategies when the confirmation strategies supply evidence that the predictions were not successful; and (5) deciding which bits of information should be held in short-term memory for further examination and which bits finally should be integrated into the reader's meaning system (K. Goodman 1970).

Although these universal strategies are constantly in operation, they take on specific characteristics depending on subject matter, reading format, organization, language style, and graphics. These variations cause the reader to apply the universal strategies in a unique way for specific materials. Dealing with only one type of written material does not help readers develop the variety of strategies they need in the multitude of reading tasks they meet in their reading lifetime.

Reading strategies are used by all readers with varying degrees of success and from the very beginning of reading acquisition. These strategies help reduce uncertainties about new materials. For reading strategies to develop and become more effective, readers must avail themselves of all the language systems which can provide them sufficient clues. It is also essential that readers deal with natural written language.

When learning takes place in context, learners are able to induce their own rule systems. They categorize certain items, notice how those items operate differently in different contextual settings, and organize the developing concepts in a manner that has been successful for them since infancy. However, when learners are taught specific rules without the opportunity to set up their own categories through contrasting and comparing in context, they are apt to become confused when a rule doesn't operate the way it was taught. For example, when the word *can* is taught out of language context, relating it to the spelling pattern of *an* and giving it the pronunciation often used in list stress like /keen/, a reader must reject this rule in order to read the sentence *Can you put it in the trash can?* successfully. The question marker in this sentence is usually pronounced /kin/ or /kan/. The two different *cans* have no grammatical or semantic similarity in the context of the sentence.

Working with natural written language, the reader relies on the semantic and syntactic system as well as a graphophonic system. The reading context is not as ambiguous as it is when isolated units of language are taught. In context, the reader can use grammar and meaning to decide on the connotative meaning of a linguistic unit and thus can approach the author's meaning more effectively. In natural language any single language unit may have a different sound, be a different part of speech, or have a different meaning from the way it was taught in isolation. To isolate units of language and teach them directly to readers confuses learning and increases complexity, for it creates a more abstract task for readers than they face with whole language.

Reading Strategy Lessons

The teacher's role in reading is significant. The teacher must discover which reading strategies the reader is using effectively and which strategies the reader needs help to develop. The teacher should also be able to find, or write if need be, material with linguistic strengths that enable readers to develop effective strategies. These written materials which provide the reader with supportive reading are called reading strategy lessons. (Goodman, Burke, and Sherman 1975). Reading strategy lessons help make learning to read a less abstract task and make it possible for readers to use their language competence to the fullest as they discover how to become independent readers.

A strategy lesson is a carefully constructed instructional plan developed for an individual reader or a group of readers. Evidence

that a reader needs the support emphasized in the lesson could come from a miscue analysis of a student's reading (Goodman and Burke 1972). Strategy lessons help readers focus on aspects of written language they are not processing effectively. Literary quality, interest to the reader, and significant context must be of prime concern even though the strategy lesson emphasizes a specific aspect of written language. Learning by using specifically focused material which is as much like "real" reading material as possible gives readers all the language cues intact so they can interrelate the specific strategy into a total language setting.

Almost all readers can benefit from reading strategy lessons: (1) proficient readers who are effective users of reading strategies; (2) readers who are developing effective reading strategies but do not use them consistently; and (3) ineffective users of reading strategies.

Proficient Readers

Proficient readers generally use reading strategies effectively without instruction. Reading strategy lessons help these proficient readers feel confident that their reading strategies are indeed appropriate. These readers should be encouraged to expand their reading experiences. Strategy lessons for this group should focus on questioning the truth of what is read and deal with subtleties and inferences not stated by the author. Such readers could be provided with folk tales that retain the uncommon or archaic language structure of the people who told the tales originally, for such structures broaden readers' awareness of the rich variety of language possible in English.

Inconsistent Readers

The second group of readers shows evidence of good use of reading strategies, but they are not consistently proficient. These readers use effective reading strategies when the material is highly interesting to them or when it is easy because it has a low concept load. However, when these readers find themselves reading material which is complex, they use less efficient reading strategies. They stop searching for meaning and end up sounding out or word calling. When asked how they handle any particular reading problem, such readers often say they sound words out; they may be unaware that they use context to read or they may believe the teacher disapproves of it. Strategy lessons help these readers become aware of the various effective reading strategies they already

use when reading easy material, permitting them to transfer effective reading strategies to more difficult reading materials.

One type of reading strategy lesson for such readers is a story with a concept or word which is probably not well known, as shown in this lesson written by the author:

> The boy was looking for Petoskies. He was walking slowly to make sure he wouldn't miss them. He usually found a number of them each time he went looking for them. They were not easy to find because they were the same color as the sand. He enjoyed looking for Petoskies on the beach. He was helping his mother, too, since she used them in her work. She was an artist and made jewelry with them. Petoskies are usually bluish gray in their natural state with the fossils in them somewhat darker. When Petoskies are polished, the gray color becomes lighter and the fossils take on a brown character. Petoskies are found only on the shores of the Great Lakes.

Put this story on an overhead projector and use it with a small group of readers. Tell them not to worry about pronouncing every word as they read. Cover the entire story and move the cover down, exposing one more sentence with each move. As each sentence is exposed ask the children to tell what the word (point to Petoskies) means. After each sentence, ask the children to revise their guesses about the word. Do not pronounce the word for the children, nor should you ask them to pronounce it. If any reader does say the word, the pronunciation should be accepted without comment. Only after the story is completely exposed and the meaning of Petoskies fully discussed should you ask for variations in pronunciation and finally tell the group how you think it may be pronounced. This is an interesting lesson because many teachers may not pronounce Petoskies the same way the people who polish and sell these stones do.

Another strategy lesson for this group of readers would help them find ways of deciding when words are significant and their meaning should be pursued, and when they are insignificant and can be omitted without losing a great deal of what the author wants to say. Such a strategy lesson can be written by the teacher or can be found in available reading material:

> All the Christmases roll down toward the two-tongued sea, like a cold and headlong moon bundling down the sky that was our street; and they stop at the rim of the ice-edged, fish-freezing waves, and I plunge my hands in the snow and bring out whatever I can find. In goes my hand into that wool-white bell-tongued ball of holidays resting at the rim of the carol-singing sea, and out comes Mrs. Prothero and the firemen (Thomas 1954, pp. 5-6).

The reading of this selection should be followed with discussion questions: Which words or phrases can you leave out and still get enough out of the passage to understand what the author wants you to know? Are some words or phrases meant to bring out feelings in you rather than for you to understand their exact meaning? How can you tell?

It is a significant insight when students realize that all readers make judgments when they read. They make judgments about what is significant and needs to be understood immediately; what may be felt and not fully understood; what can be actually omitted; or what can be not understood at all but kept under wraps for another time when the same concept reappears in a richer context. Readers must also understand that they are the decision-makers at each of these steps and no one else can make these decisions for them if they are to become independent readers.

Another strategy lesson makes use of material in which certain words or phrases unknown to the reader are eventually explained in the context. The reader must read on so the author's use of appropriate redundancy in written language can provide the reader with additional meaning cues. Again, the teacher can write such a story or find one in ready-made materials:

> In the high and far-off times the elephant, o best beloved, had no trunk. He had only a blackish, bulgy nose as big as a boot that he could wriggle about from side to side; but he couldn't pick up things with it. But there was one elephant—a new elephant—who was full of "satiable curtiosity," and that means he asked ever so many questions (Kipling 1965, p. 63).

Inefficient Readers

The third group of readers needs help in realizing that the purpose of reading is to gain meaning for one's self and not to complete a task in a way specified by a teacher for an unknown purpose (Watson 1973). For a variety of reasons the development of effective reading strategies is disrupted for such readers. They use effective reading strategies occasionally in short phrases or sections of written material, but in most reading situations these readers tend to omit words they think they do not know; they do not predict acceptable grammatical or semantic structures as they read; they read word for word using sounding out techniques without concern for meaning. They do everything they were taught to do in an isolated and unrelated fashion. They look for little words in big words and find *fat/her* an acceptable solution for *father*. They separate words between two middle consonants and

often read *lit/tul* for *little* and *prit/tee* for *pretty*. When they do make occasional effective use of reading strategies they lack confidence in deciding which strategy is most effective. They regress and correct in situations when it is inefficient to do so. For example, if such a reader reads *can't* for *cannot* because of the use of an appropriate predicting strategy, this is corrected when the reader picks up additional graphic cues. Such readers often think that graphic input is the most significant aspect of reading. Reading is not to discover something new or for enjoyment, it is to satisfy another person.

The first step for such readers is to realize that they are effective users of language and that being a language user can help in reading. Using a modified cloze procedure as a strategy lesson may help such readers become confident about their ability to use context to predict and guess at missing units:

> "Help!" said Andy. "Help! Help!"
>
> Dan and Bill came fast. They saw Carlos down in the water.
>
> "Watch out!" said Dan. "That _____ is after you!"
>
> Carlos splashed and splashed the water at the _____ to get it to go away. The _____ did not like the water splashed at it. It swam under the water away from Carlos. Then Carlos swam fast through the water to the boat.
>
> "Fast!" said Dan. "Help pull Carlos back up on the boat. The _____ is coming back!"
>
> The men helped Carlos out of the water.
>
> "Are you all right?" said Dan.
>
> "I am all right," said Carlos. "But that _____ was after me. It was a good thing it went away when I splashed water at it (Berres et al. 1959, pp. 24-25)."

Selecting the grammatical item which will be missing should be done carefully for these readers. They should be given only highly predictable blank slots in the first few strategy lessons, since the most important first step is to build self-confidence.

Because these readers concentrate on the relationship between letters and sounds and minimize the use of grammatical and semantic information, they often develop habitual associations between words or phrases that have close graphic and phonemic similarity. Examples of such habitual associations include *for* for *from*, *saw* for *was*, *through* for *thought*, and *this is* for *is this*. The best way to help readers break down habitual associations is to prepare strategy lessons that carefully control the linguistic

setting so that only one of the two habitually associated items can fit the meaning. The material must be within the experiential understanding of the readers so that the error is obvious if they read the inappropriate term.

The developmental sequence of a strategy lesson plan which would break down habitual associations includes: (1) Write an experience story to elicit the particular words the child or children should think about. For example, if a reader or small group in a class habitually associates *was* and *saw*, the teacher might suggest writing an experience story about the tour of the school the students took the previous day, emphasizing what they *saw* on the trip *yesterday*. This would cause the students to use the word *saw* frequently during dictation and would cue them to read the word *saw* when they read the experience story. (2) Find or write a story with only one of the habitually associated words used frequently in an unambiguous language situation. A few sentences from such a story might be: "John looked up and down the street. He wanted to tell his father about the good news. When would his father come? Finally, as he looked down the street, he saw Father coming." In this setting, if the reader is focusing on understanding, the clause *he saw Father* could not possibly be read *he was Father*. (3) When the reader or readers show evidence of handling one of the habitually associated words, write or find some material where the other habitually associated word is used in grammatical settings whose context builds toward the appropriate word so the confusion will be less likely to occur. (4) Encourage the students to read in regular situations to see if they have developed the effective reading strategies needed to handle the habitual association.

In summary, strategy lessons help readers develop appropriate reading strategies in a context of unambiguous language so readers have the support of meaning and grammatical systems to help them predict and confirm. Teachers are the key to the instruction program suggested here. They need to know what kinds of support readers require and how to select and write material to best help develop effective reading strategies. They must supply readers with appropriate language settings through which the readers can induce their own rule systems. The need for teachers knowledgeable about language, how language is learned, and what the reading process is like, becomes crucial. Such a teacher is the single most important person helping students learn to read effectively.

REFERENCES

Berres, F.; J. Coleman; W. Brisco; I. Hewett. *The Sea Hunt*. San Francisco. Harr Wagner Publishing Company, 1959.

Goodman, Kenneth S. "Behind the Eye: What Happens in Reading." In *Reading: Process and Program*, edited by Kenneth S. Goodman and Olive Niles. Urbana, Illinois: National Council of Teachers of English, 1970, 3-33.

Goodman, Yetta M., and Carolyn Burke. *Reading Miscue Inventory: Manual*. New York: Macmillan Publishing Co., Inc., 1972.

Goodman, Yetta M.; Carolyn Burke; Barry Sherman. *Strategies in Reading*. New York: Macmillan Publishing Co., Inc., 1975.

Kipling, Sir Rudyard. *Just So Stories for Little Children*. New York: Schocken Books, 1965.

Thomas, Dylan. *A Child's Christmas in Wales*. New York: New Directions, 1954.

Watson, Dorothy J. "Helping the Reader: From Miscue Analysis to Strategy Lessons." In *Miscue Analysis: Applications to Reading Instruction*, edited by Kenneth S. Goodman. Urbana, Illinois: National Council of Teachers of English and ERIC Clearinghouse on Reading and Communication Skills, 1973, 103-115.

3 Miscue Patterns
The Relationship of Beginning Reading Instruction and Miscue Patterns

Jayne A. DeLawter

Miscue research is the focus for an increasing number of studies in reading. A miscue is an unexpected oral reading response (K. Goodman 1965). Proponents of miscue research assert that the study of miscue patterns leads to greater understanding of the reading process and, in turn, to improvement of reading instruction for children (DeLawter 1970; K. Goodman 1969; K. Goodman and Burke 1970; Y. Goodman 1970).

Until now, however, most miscue research paid little attention to school children's previous instructional experiences. The usefulness of this research is consequently limited unless possible sources of miscues are more thoroughly investigated. The interrelationship among formal instruction, experience, and miscue patterns warrants attention in order to clarify miscue findings and to provide guidance in instructional planning.

For example, at the present stage of miscue research, it is difficult to determine whether a miscue pattern demonstrates a lack of exposure to or minimal competency with a particular strategy, an overlearned or overused strategy based on previous instruction, or a stage in the development of reading competence. Planning appropriate instruction at any level of competence would be greatly facilitated by knowledge about the sources of miscues.

In addition, a study of the interrelationship of miscue patterns and instructional methodology might provide insight into general school performance, i.e., the ways in which strategies and attitudes children acquire in their central primary school experience—reading—are carried over into other tasks and areas.

Barr, in a study of the relationship of word recognition errors and instruction, finds evidence that instruction does influence reading strategies in predictable ways. As she acknowledges, though, her study was limited to words tested in isolation rather than context, and her data were collected after only a brief period of instruction.

She suggests the need for further research which assesses "the effect of different instruction on children's reading strategies over longer time periods using contextual materials (1972, p. 527)."

Independently of Barr's work, the present study was undertaken to determine if a relationship between beginning reading instruction and patterns of miscues could be shown to exist over an extended period of time.

Description of the Study

One hundred sixty-nine children were involved in the study. All were second graders in New York City public schools located in low income areas. The children had participated in a large research project* for two and one-half years and had received specific reading instruction in first and second grade in one of two reading approaches.

One reading approach incorporated "linguistic materials with an emphasis on decoding and no emphasis on meaning of any kind (Goldberg and Taylor 1969, p. 1)." This program was derived from the Allen and Allen *Read Along With Me* materials which served primarily to familiarize children with the regular spelling patterns of English.

In classroom exercises, nonsense responses were acceptable when they indicated a child's use of appropriate sound-symbol relationships. The content of the stories was less important than the practice the stories offered in decoding. Two reading series which controlled vocabulary on the basis of regular spelling patterns were used: the *Miami Linguistic Readers* and the *Merrill Linguistic Readers*.

The second reading program followed a sight word approach, incorporated in a basal reader series, the *Chandler Language Experience Readers*. These materials emphasized multi-ethnic characters and stories cast in an urban environment. They sought to use vocabulary and sentence patterns characteristic of children's natural oral expression. These were derived, in some cases, from transcriptions of actual recorded conversations.

Photographs of city life were included and boys played important roles in many of the stories. Typical of basal series, the content of the stories was heavily emphasized, pictures were primary clues to meaning, and word identification skills were largely peripheral in the presentation of the stories.

*The Beginning Reading Project, directed by Dr. Miriam Goldberg, was sponsored by the Center for Urban Education and Teachers College, Columbia University.

Of the seven reading approaches used in the Beginning Reading Project, these two programs (hereafter, "decoding emphasis" and "meaning oriented") were considered to be at opposite ends of the continuum with respect to the importance of meaningful content.

A check was made on instructional methods used by the second grade teachers. Supervisors in each school were asked to match participating teachers with one of two prepared lessons. The lesson selected was to be typical of the teacher's usual reading lesson in content and procedures. There was only one mismatch in the group of 11 teachers.

For this study the children read 30 test words presented in isolation or in context. These words, all of more than one syllable, did not appear in the formal instructional program of either of the reading approaches for first or second grade mastery.

In the context presentation, the test words, used as nouns, were embedded individually in sentences which provided relevant clues to the meaning of the words but did not limit logical responses to those words. Substitution of other words could result in meaningful but inaccurate sentences.

Table 1

Percentage of Nonword and Real Word Miscues by Instructional Group

Miscue	Decoding Group		Meaning Group	
	Isolation and Context	Context Only	Isolation and Context	Context Only
Nonword	70	65	41	46
Real word	30	35	59	54
Total	100	100	100	100

Table 2

Percentage of Miscues According to Quality of Decoding by Instructional Group

Quality of Decoding Attempts	Decoding Group	Meaning Group
Good	31	15
Fair	31	21
Poor	38	64
Total	100	100

Each child was tested for about 15 minutes on either the isolation or context presentation. Sessions were tape recorded and responses were also noted on a record sheet by the examiner.

Miscues were classified as substitutions or omissions. Hypotheses were tested with chi square tests of independence. Additional analyses were performed on the substitution errors from a random sample of 24 protocols from each instructional group. Frequency of real word miscues, quality of decoding attempts, and relevance of real word miscues to their context were evaluated.

Results

Patterns of Miscues. Miscues, both words and nonwords, were made by children from both instructional groups. However, the groups demonstrated different patterns of miscues. The decoding emphasis group offered about twice as many nonwords as words, while the meaning emphasis group had a higher percentage of real word miscues than nonwords. These trends were present in both types of presentation. (See table 1.)

When miscues were evaluated with respect to the quality of decoding,* children who had been taught by the decoding emphasis demonstrated over 60 percent good and fair attempts. The majority of miscues (64 percent) made by the meaning emphasis group were classified as poor attempts. (See table 2.)

Several individuals demonstrated patterns of miscues which indicated overuse of strategies stressed in their reading lessons. For example, the decoding emphasis approach suggested instructional activities in which children read lists of rhyming words. New words were introduced in this way and nonsense forms were frequently included on the lists to check for decoding ability.

When asked to read the words presented in isolation, several children taught by this method read the test words (none of which rhymed with each other) as if they were rhyming words. The following responses made by two children illustrate this occurrence:

shoulder ---------- spook	poison ---------- pon
superintendent ---------- rink	correction ---------- ron
pilot ---------- pink	secretary ---------- son
station ---------- skook	stadium ---------- stom

*Categories for evaluating decoding attempts were *good* (use of graphophonic elements and syllabication similar to the test word), *fair* (use of a few graphophonic elements), *poor* (use of initial sound only; response with little or no apparent graphophonic resemblance to the test word).

Apparently these children noted some of the graphic features and responded to the task in a way that was similar to their classroom reading experiences.

Patterns of miscues among children taught by the meaning emphasis approach did not as dramatically illustrate specific strategies. These children did tend to offer hasty responses, while children in the decoding emphasis group frequently took longer attempting to figure out unknown words.

Syntactic Acceptability. In the context presentation for both instructional groups, most of the miscues that were real words fit into the syntax of the sentences. In fact, even some of the miscues which were judged syntactically unacceptable fit with the preceding syntactic context, though they were not acceptable in the total sentence. For example, in number 5, the word *sensible* is not sufficient to complete the sentence. However, it is not a completely inappropriate response, since it would be acceptable if followed by a noun.

No patterns emerged which differentiated the groups on the basis of syntactically acceptable miscues. The following examples illustrate word miscues made by children from both groups (miscues are italicized and test words are capitalized):

Acceptable

1. The boy had to wait for a *chance* (MINUTE) before he could get on the bus.
2. The man standing by the plane is a *policeman* (PILOT).
3. The clothes were very dirty, so Mother used *deodorant* (DETERGENT) to wash them.
4. Mr. Stone didn't shave for three weeks. Now he has a real *noisy teacher* (MOUSTACHE).

Unacceptable

5. The teacher told Tom to take the note to the office and give it to the *sensible* (SECRETARY).
6. Mike is a boy all the children like. He doesn't have any *empty* (ENEMY) in the class.

Semantic Acceptability. For both groups, about half of the miscues that were real words were judged relevant to the meaning of their context. As indicated earlier, the majority of miscues made by the meaning oriented group were real words, while the majority of miscues made by the decoding emphasis group were nonwords.

Therefore, the actual number of miscues which were semantically acceptable was higher for the meaning oriented group.

The following sentences include examples of semantically acceptable miscues:

1. That costs too much. I only have a *dollar* (QUARTER).
2. A window in our apartment was broken, so we called the *window fixer* (SUPERINTENDENT).
3. When Ann saw a rat, her mother said, "We need some *mouse traps* (POISON)."
4. The teacher told Tom to take the note to the office and give it to the *principal* (SECRETARY).
5. Since it was raining, the game couldn't be played in the *sunlight* (STADIUM).
6. Nobody was hurt, but a lot of windows were broken during the *fire* (RIOT).
7. Mr. Stone didn't shave for three weeks. Now he has a real *beard* (MOUSTACHE).

Some examples of miscues which were judged unacceptable:

1. You can trust her. She won't break a *problem* (PROMISE).
2. Ted rode his *birthday* (BICYCLE) on the path.
3. While the man was taking pictures, he dropped his *camp* (CAMERA).
4. When they were married, they had a big *window* (WEDDING).
5. The shop was burning! Firemen rushed to the *rubber* (RESCUE).

Summary. The majority of miscues made by the decoding emphasis group were nonwords which resembled test words closely in graphic/sound elements. Almost all of this group's real word miscues were syntactically acceptable and about half of them were semantically acceptable.

The majority of miscues made by the meaning oriented group were real words which were not especially close to the graphic/sound elements of the test words. As with the other group, almost all real word miscues were syntactically acceptable and about half were semantically acceptable.

Limitations

This study, as an outgrowth of the larger Beginning Reading Project, was limited by the complexities inherent in longitudinal

research efforts in contemporary urban schools. Despite regular project teacher training sessions, classroom visits, and formal observations, it is difficult to state with certainty the degree to which the reading programs were carried out as intended.

Discussion

This study sought to answer these questions: are there differences in miscue patterns among children taught by different beginning reading approaches? What is the nature of these miscues?

The data from the study suggest that there are miscue patterns which appear to be directly related to beginning reading instructional approaches. These patterns are consistent with the different emphasis of each instructional program and demonstrate strategies that are predictable, given particular instructional emphases.

Significantly, the miscue patterns distinguish the two groups after two years of reading instruction. A child's experiences over a period of time might be expected to lessen the recognizable effects of specific reading instruction because he or she has had time to personalize and modify the strategies. The fact that the patterns are still distinguishable in spite of these other influences reinforces the findings of the study.

The miscue patterns of the two groups differed most strikingly with regard to the frequency of miscues (1) which were real words and (2) which were similar to the graphic/sound elements in the test words. Children taught by a meaning oriented approach tended toward miscues which were real words, while those with an emphasis on decoding produced mostly nonword miscues.

The fact that real word miscues offered by both groups fit the context only about half the time has clear instructional implications. It suggests the need for instruction which helps children realize that they are not "reading" if the words they are saying aloud or thinking do not make sense in the context. Indeed, only as far as they are aware of the meaning of the passage will individual words be significant.

There are few reading approaches which suggest actual systematic strategies for "reading for meaning." Although this phrase has become a popular cliché, its meaning is still rather vague. Good teachers may ask children if a certain word makes sense in a passage, as did several teachers in the meaning oriented approach in this study. They may also ask for a retelling of the selection to evaluate the extent of the child's understanding of it. However, because reading is commonly seen as an accumulation of skills and because "meaning" is such a personal and mysterious concep-

tual experience, few real strategies other than these have yet been devised to help children reconstruct meaning from printed materials.

The data from this study show that children from both instructional groups made few real word miscues that interfered with or changed syntax. This finding may be a function of the syntax provided in the test situation. It might also be related to the nature of syntactic systems.

In this study the test words, used as nouns, appeared near or at the end of individual sentences. These statements were not related to each other in any meaningful way. Children may have gotten used to this grammatical pattern as they read the items. If so, implications regarding syntactic patterns would require further study.

Another explanation might be that the two instructional approaches were neutral about syntax. It is generally accepted that children acquire control over grammatical conventions before they begin formal reading instruction. Perhaps the conventions are so firmly established that instruction in reading of the type reported here does not affect them.

The strength of the syntactic component should be viewed as a bonus for teachers. Grammar, a highly constraining system, can be used actively as a strategy for generating meaning. Children should be helped to use more consciously their already developed language abilities as they read.

The finding that patterns of miscues relate to instructional approaches has important implications for diagnosis. If particular patterns of miscues are noted for a specific child, the diagnostician should immediately check the child's previous instruction. If the patterns are consistent with the instructional approach, the diagnostician will have valuable information. The patterns suggest that the child has indeed learned reading strategies, but for some reason is unable or unwilling to use them appropriately. This knowledge would guide future instructional planning. For example, the diagnostician could then identify new strategies and new goals which do not duplicate previous activities.

If the child's pattern of miscues is not predictable from previous instructional experiences, different kinds of questions would need to be asked. For instance, did the child receive instruction in a particular strategy? Was the child absent when key strategies were presented? What is the child's preferred sensory modality? What are the characteristics of the instruction that apparently was not learned?

The ways in which findings from this study can be used depend, in part, on children's stages of development. Some miscues are more reasonable at one stage than another. There is a need for further study to clarify the relationship between cognitive development, instructional experiences, and patterns of miscues.

Other areas for further research are also suggested. A similar investigation using longer, more varied reading selections might be conducted with young children. A study with older children might determine whether strategies taught in their beginning reading programs were still evident.

Another study might look at the degree to which children generalize the strategies they learn in specific areas like reading. The present findings suggest that learned strategies are persistent; it is reasonable to ask how they affect other curricular areas and ways of approaching problems. For instance, is a meaning oriented approach in beginning reading more likely to encourage critical thinking and personal interest in reading for pleasure? Does a decoding emphasis in early instruction lead to highly organized and precise ways of approaching problem situations? Findings such as the ones discussed could bring perspective to the rather limited current emphases in reading instruction.

REFERENCES

Allen, Robert A., and Virginia F. Allen. *Read Along With Me*. New York: Bureau of Publications, Teachers College, Columbia University, 1964.

Barr, Rebecca C. "The Influence of Instructional Condition on Word Recognition Errors." *Reading Research Quarterly* 7 (1972): 509-29.

Carrillo, Lawrence. *Chandler Language Experience Readers*. San Francisco: Chandler Publishing Company, 1966.

DeLawter, Jayne A. "Oral Reading Errors of Second Grade Children Exposed to Two Different Reading Approaches." Ph.D. dissertation, Teachers College, Columbia University, 1970.

Fries, Charles, et al. *Merrill Linguistic Readers*. Columbus, Ohio: Charles E. Merrill Books, Inc., 1966.

Goldberg, Miriam L., and Marian S. Taylor. "Working with the Urban Disadvantaged: Beginning Reading Project." New York: Center for Urban Education, Teachers College, Columbia University, May 1, 1969. (Mimeographed.)

Goodman, Kenneth S. "A Linguistic Study of Cues and Miscues in Reading." *Elementary English* 42 (1965): 639-43.

——. "Analysis of Oral Reading Miscues: Applied Psycholinguistics." *Reading Research Quarterly* 5/1 (Fall 1969): 9-30.

Goodman, Kenneth S., and Carolyn L. Burke. "When a Child Reads: A Psycholinguistic Analysis." *Elementary English* 7 (January 1970): 121-29.

Goodman, Yetta. "Using Children's Miscues for Teaching Reading Strategies." *Reading Teacher* 23 (February 1970): 455-59.

Miami Linguistic Readers. Experimental edition. Boston: D. C. Heath and Company, 1965.

A Different Look at Reading in the Content Areas

Kenneth L. Carlson

School reading programs have developed on a twofold premise. First the child learns to read, and then the child reads to learn. It is assumed that reading instruction enables children to develop the ability to read in a variety of content areas.

Educators, however, report that many children seem to be able to read and comprehend grade level material in their basal readers, but they appear to lack the skills to read and comprehend subject matter texts for the same grade level. Many educators suggest that the readability of content area materials requires closer scrutiny because it may actually be higher than publishers designate. Some educators question the conceptual load of such materials and suggest that too many ideas new to the reader may be involved. All seem to agree that many elementary and secondary students have difficulty reading content area material.

Textbooks on reading instruction seldom cite research as a basis for the suggestions offered teachers who deal with reading in the content areas. Durrell (1956) emphasizes the idea that the critical thinking ability of the reader determines how well he or she reads subject matter. He considers critical thinking abilities important for reading in the content areas. The reader must be able to select material pertinent to a topic or particular occasion. He or she must be able to find exception with the author's viewpoint, distinguish sense from nonsense, sort fact from opinion. Critical reading includes finding differences in viewpoints, noting the author's bias, and locating exaggerated statements and unfounded claims (pp. 305–307).

Some Old Studies

Gray (1960) reports "a range of correlation coefficients from .22 to .70" in early studies comparing reading achievement and scholastic standing (p. 1126). Bond and Tinker (1957) report that "the

correlations between general reading tests and reading tests in the content fields range from about .30 to .50 (p. 352)." Traxler (1957), however, states:

> The existence of significant positive correlation between reading ability and achievement in a subject field does not necessarily mean that improvement in reading ability will result in improved achievement (p. 9).

Summers (1957) reports that recent studies of reading in the content areas used "unique or more sophisticated measuring techniques (pp. 137–138)." He summarizes a study where one researcher used an introspective technique to study how students viewed reading difficulties in subject matter areas. Cafone's doctoral research (1966) also focuses on asking students to communicate information about their reading process as they read literature and science material. Smith and Dechant (1961) emphasize that a variety of reading skills are required in subject matter reading.

A survey of reading research in the content areas reveals that the majority focus on such factors as paragraph meaning, readability factors, growth in reading and subject matter achievement, and the effects of varied instructional methods on comprehension in content subjects. The research reported in the literature is concerned with measuring the results of the reading process after reading subject material rather than studying the reading process as it is applied by individual readers (Traxler 1941; Traxler and Townsend 1946, 1955; Robinson 1962–65; Weintraub, Robinson, and Smith 1965–70; Weintraub et al. 1971–73).

Two New Studies

In 1970 this writer conducted a study of reading in the content areas that focused on analyzing the process that readers used as they read various subject matter (Carlson 1970). A small group of fourth grade students orally read two basal reader stories, two social studies selections, and two science selections, all taken from materials new to the readers. The oral readings were tape recorded to capture the oral reading miscues for analysis. At the end of each selection the child was asked to recall and retell the selection orally, permitting an analysis of the reader's comprehension. This, too, was recorded.

The reading miscues were analyzed using the *Goodman Taxonomy of Reading Miscues* (K. Goodman 1969). This taxonomy is a psycholinguistic research instrument. It includes categories such as phonemic and graphemic relationships, morphemic considerations, grammatical function, and syntactic and semantic accept-

ability. Miscue patterns in the various content selections were analyzed and compared.

Kolczynski (1973) completed a similar research study. He analyzed the readers' use of syntactic and semantic cueing systems in their effort to gain meaning from literature, science, social studies, and mathematics passages. He used the *Reading Miscue Inventory* (Y. Goodman and Burke 1972), which is a shorter version of the *Goodman Taxonomy of Reading Miscues*.

Both research studies assumed that reading is not just a process of matching sound with symbol, but involves the interaction of the reader with the language of the printed page. The reader uses his or her memory bank of conceptual experience and language knowledge in an attempt to comprehend the written message encoded by the author. As the reader interacts with the printed page, he or she predicts and anticipates what language structures will be used to transmit the intended message. A reader becomes increasingly proficient as he or she increases in ability to reconstruct meaning.

Reconstruction of meaning involves reading the signals of at least three distinct but interrelated cue systems. The reader is cued by graphophonic, syntactic, and semantic information. If the reader successfully processes this information and correctly anticipates the language of the page, no deviation will be noted. However, when the reader misprocesses or the prediction deviates from the cues on the page, he or she produces a miscue. A miscue is defined as any deviation from the expected response to the printed page. These miscue phenomena form the basis for these two studies, which provide some additional insight into the complexities of the reading process and give us a perspective that permits a different look at reading in the content areas.

Summary of the Findings

The readers in the Carlson and Kolczynski studies exhibit the greatest similarity when reading basal reader or literature selections, which may be a result of the instruction they received. Analyzing the miscues reveals several patterns. The majority of miscues in all the reading selections show a high degree of partial or full syntactic acceptability and a medium or high graphophonic similarity. A large number of miscues are totally syntactically acceptable and show slight or no syntactic change. This indicates that the subjects used syntactic information to cue them in their reading.

There is a tendency for miscues with partial or full semantic acceptability to have medium or high graphophonic similarity.

Subjects tend not to self-correct miscues that are partially or totally semantically acceptable. There is a strong tendency for miscues to have the same grammatical function as the expected text response.

Although subjects exhibit the greatest similarity on basal reading selections, it is apparent that the subjects' ability to read content area materials cannot be easily predicted from their performance on basal reader material. There is little relationship between the subjects' miscues per hundred words and their comprehension of the selections. All subjects tended to use graphophonic cues to aid them with all types of materials; similarly, these subjects made extensive use of syntactic and semantic cues to gain meaning, regardless of content area.

Previous writings on reading in the content areas have assumed that readers need to shift their modus operandi when reading different subject matter. These studies suggest that the reading process remains relatively stable across passages from various content areas.

Implications

For the most part, the conclusions drawn from the data in the Carlson and Kolczynski studies closely parallel the results of related research using the *Goodman Taxonomy of Reading Miscues* and the *Reading Miscue Inventory* as a basis for data analysis (K. Goodman and Burke 1968, 1969, 1973). The assumption that reading is a psycholinguistic process is apparently further substantiated by the results of these studies.

Reading is a language process to which readers contribute information from their knowledge of language and their experience. When this information closely matches the language usage and experiences of the writer, a meaningful reconstruction of the intended message takes place. If this description is accurate, the present methodology of so-called developmental reading instruction needs critical examination. Much time and energy is now being devoted to reading instruction in basal reader programs. It may be that the rigid structure in such programs does not lead to as broad a development of reading behavior as educators once believed.

It has been hypothesized that the careful regimented pattern of basal reading instruction will eventually lead to development of a broad base of reading skills that will enable the reader to process a variety of contexts and grammatical structures. Indeed, this may not be the case at all. Perhaps the current methods of basal reading

instruction, when followed too rigidly, are rather a type of controlled retardation. In any case, further examination of current practices and the results of present-day reading instruction is required.

One of the practices that must be examined further is the word analysis program in developmental reading instruction. Most of these programs consist mainly of phonetic and structural analysis methods of word identification; almost all basal reading programs have a high degree of vocabulary control. The graphophonic cue system is useful in reading, and mature readers exhibit considerable mastery in using this cue system as they read. However, when phonics instruction and development of sight vocabulary are isolated elements in reading instruction, pupils may consider them separate rather than related elements.

An overemphasis on phonics instruction and too great a concern for mastery of the controlled vocabulary in the basal reading program can limit, rather than extend, the reading skills of youngsters. The exclusive use of the graphophonic cue system may encourage maturing readers to rely on a method of word identification that probably can only be partially successful at best. These content area studies seem to indicate that the readers resorted to a graphophonic approach when the reading became difficult for them: to some extent they seemed to follow the rule, "When in doubt, sound it out." These readers also showed a strong reliance on syntax to cue them as they read.

Of course, sight vocabulary development is necessary if a youngster is going to develop into a mature reader, but too often the approach is more like exposure to a printed symbol matched by the utterance of the verbal label for that symbol. Thus, children are shown *t-o-b-a-c-c-o* and told that the word is *tobacco*. All too often, the instruction does not include a variety of strategies that focus on a fuller understanding of the *concept* of tobacco. Unless the conceptual aspect of vocabulary development receives attention, students may be left to their own devices to assimilate the semantic input needed to read content area material.

Recommendations

Several recommendations can be made to classroom teachers and curriculum directors. There should be careful analysis of content area materials used for instruction. It is insufficient guarantee that publishers designate materials as suitable for students in a certain grade level; such materials must be tried out on a

pilot basis to see if children can handle the syntactic and semantic information with relative ease. Children appear to rely on syntactic cues to aid them in reading as much as they rely on graphophonic information. Teachers should be aware of this and provide instructional strategies that encourage children to make effective use of such cues as they read.

Considerable time and energy is spent planning and conducting directed reading activities with basal reading materials in upper elementary classrooms. It may be wiser to reallocate at least some of this time and energy to reading instruction with content area materials. Particular attention should be given to helping children understand the concepts presented in such materials. Teachers should observe and analyze children's miscues on a regular basis; the tape recorder provides an effective and economical way of recording children's reading so teachers can analyze it later. Selections should include samples from content area books, for such analysis can provide the teacher with information about the children's strengths and weaknesses that will permit appropriate individualization of instruction.

Instruction in content area subjects cannot be limited to a single textbook. A wide variety of reading materials should be used to meet the reading needs of children with different interests and varied experiences. Teachers should use real life experiences or provide vicarious experiences with films, pictures, and simulation activities in conjunction with oral discussion to allow children to exchange ideas. Such activities are necessary to develop fully the concepts presented in text material.

Conclusion

The so-called reading to learn phase of reading development does not seem to be exhibited very well by the subjects in these content area studies. Serious consideration must be given to current approaches in content area instruction as well as reading instruction. Guidelines should be developed for analyzing, selecting, and producing content area material for instructional purposes.

Reading is *simultaneously* a sensory process, a perceptual process, a linguistic process, and a cognitive process. Thus, the reading process is psycholinguistic by nature. Further research investigations must not ignore the simultaneity of these elements. To investigate only the sensory or perceptual aspects of the reading process without some concern for the linguistic and cognitive

aspects may only perpetuate some of the simplistic notions that already permeate the research literature. Reading is indeed very complex and only by researching the entire process will researchers gain valuable insights.

REFERENCES

Bond, Guy L., and Miles A. Tinker. *Reading Difficulties: Their Diagnosis and Correction*. New York: Appleton-Century-Crofts, Inc., 1957.

Cafone, Harold. "Individual Differences in the Reading Process of Ninth Grade Retarded Readers." Ph.D. dissertation, University of Arizona, 1966.

Carlson, Kenneth L. "A Psycholinguistic Description of Selected Fourth Grade Children Reading a Variety of Contextual Material." Ph.D. dissertation, Wayne State University, 1970.

Durrell, Donald. *Improving Reading Instruction*. New York: World Book Company, 1956.

Goodman, Kenneth S. "Analysis of Oral Reading Miscues: Applied Psycholinguistics." *Reading Research Quarterly* 5/1 (Fall 1969): 9-30.

Goodman, Kenneth S., and Carolyn L. Burke. *Study of Children's Behavior while Reading Orally*. USOE Final Report, Project No. S425, Contract No. OE-6-10-136, U.S. Department of Health, Education, and Welfare, June, 1968.

———. *A Study of Oral Reading Miscues that Result in Grammatical Retransformations*. USOE Final Report, Project No. 7-E-219, Contract No. OEG-0-8-070219-2806 (010), U.S. Department of Health, Education, and Welfare, June, 1969.

———. *Theoretically Based Studies of Patterns of Miscues in Oral Reading Performance*. USOE Final Report, Project No. 90375, Grant No. OEG-0-9-320275-4269, U.S. Department of Health, Education, and Welfare, March, 1973.

Goodman, Yetta M., and Carolyn L. Burke. *Reading Miscue Inventory*. New York: Macmillan Publishing Company, Inc., 1972.

Gray, William S. "Reading." In *Encyclopedia of Educational Research*, edited by Chester W. Harris. New York: Macmillan Publishing Company, Inc., 1960.

Kolczynski, Richard G. "A Psycholinguistic Analysis of Oral Reading Miscues in Selected Passages from Science, Social Studies, Mathematics, and Literature." Ph.D. dissertation, The Ohio State University, 1973.

Robinson, Helen. "Summary of Investigations Relating to Reading." *The Reading Teacher* XV (January 1962); XVI (January 1963); XVII (February 1964); XVIII with Samuel Weintraub and Carol Hostetter (February 1965).

Smith, Henry P., and Emerald V. Dechant. *Psychology in Teaching Reading*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1961.

Summers, Edward G. "Reading in the Secondary School." *Review of Educational Research* 37 (April 1957), 134-51.

Traxler, Arthur E. *Ten Years of Research in Reading*. New York: Educational Records Bureau, 1941.

———. "What Does Research Suggest about Ways to Improve Reading Instruction?" *Improving Reading in the Junior High School*. Washington: U.S. Department of Health, Education, and Welfare, Office of Education, Bulletin 10, 1957: 5-15.

Traxler, Arthur E., and Agatha Townsend. *Another Five Years of Research in Reading*. New York: Educational Records Bureau, 1946.

———. *Eight More Years of Research in Reading*. New York: Educational Records Bureau, 1955.

Weintraub, Samuel; Helen M. Robinson; and Helen K. Smith. "Summary of Investigations Relating to Reading." *Reading Research Quarterly*, I (Winter 1965); II (Winter 1966–67); III (Winter 1968); IV (Winter 1969); V (Winter 1970).

Weintraub, Samuel; Helen M. Robinson; Helen K. Smith; and Gus P. Plessas. "Summary of Investigations Relating to Reading." *Reading Research Quarterly*, VI (Winter 1971); VII (Winter 1972); VIII (Spring 1973).

4 The Cloze Procedure
Literacy in the Classroom*

John R. Bormuth

Children in school depend on their reading skills for academic survival. In the primary grades they read mainly to practice the reading skills they are learning. But as they enter the upper grades, their curriculum changes, dealing less and less with the basic reading skills and more and more with the serious business of learning how to live in the world they were born into. We have long believed that reading ought to be the most economical and efficient way for children to learn that content. Consequently, from about grade four onward, students are expected to acquire ever increasing amounts of knowledge and are expected to acquire that knowledge in ever increasing proportions by reading and comprehending written instructional materials.

However, when we planned things this way we may have miscalculated by assuming that students would be able to comprehend the instructional materials provided for them. In one study, I have found that 65 percent of the students in the upper elementary grades gained little or no information from the average textbook used in their instruction. In the junior high and high school the situation improved, but the proportion was still 40 percent. A recent city-wide survey in Madison, Wisconsin, obtained similar results. These data show what a lot of teachers have known for a long time—that textbooks are often so difficult many children learn little or nothing from them.

And that is what this paper is about: how a teacher can improve the effectiveness of the written instructional materials in the class-

*A more detailed version of "Literacy in the Classroom" was presented at the summer 1972 reading conference entitled "What Kids Do in Reading" at the University of Chicago. With the author's permission, it is presented here in a shortened, adapted form because of space constraints of this volume. Unintentional changes in the author's meaning are the responsibility of the volume editor and the editor of the NCRE Bulletin Series.

room. I will speak of this as "making children literate on their materials," using the term literate to mean a satisfactory level of comprehension performance. Let us postpone defining "a satisfactory level of comprehension performance" until we deal with this crucial issue in some detail later. First, let us examine some of the general aspects of literacy, as we will use that term.

Nature of Literacy

The term literacy has nearly as many meanings as it has people who use it. Some use it to refer to people who can both read and write; I will use it to refer only to people who can read. In reading, some use it to refer to people who can merely call the words that appear in print, while others use it to refer only to people who can both call the words and understand the content of printed materials. Both definitions have their virtues, but for our purposes it makes sense to speak of a person as being literate only when he or she is able to comprehend the content of written materials. The schools teach reading skills not just to make the students capable of calling words, but to enable them to get content from the printed page. Our definition of literacy follows a similar logic. Literacy is not an absolute or ideal state—people normally mean something much more sensible when they use the term. Many find their reading skills adequate to serve all important needs and feel justified in labeling themselves literate. Almost none of us claims that we have achieved absolute literacy in the sense that we can perfectly comprehend every material printed in the mother tongue. The term literate, in the common sense, is specific to the task and to the person. A person may be literate or illiterate on specific reading tasks—illiterate on this editorial but literate on that receipt, for example. We apply the label "literate" to people only when they are literate on the reading tasks they want to perform and must perform as they go about the business of achieving their life goals.

When teachers think of influencing students' literacy, they normally think of *increasing the students' reading skills.* Materials present a number of linguistic problems: deciding what each word means, deciding how those meanings should be combined in each of the grammatical patterns used by the author, and so on. Teachers normally see each of these problems as requiring some reading skill and set out to teach those skills. Editors or writers influence students' literacy by *altering the readability of the materials.* They assume readers know certain skills and select language that will present their information without requiring other linguistic skills.

We can influence students' literacy on a reading selection in either of two ways: teach them more of the skills required by the selection or alter the materials.

Teachers typically cannot rewrite textbooks. But they can assess the readability of materials and select those most suitable for their students, and they can prepare a child for a reading assignment to reduce the difficulty of the material.

A teacher's basic problem is to match materials with students. Ideally, every material sold for use in schools should tell how much skill is required to read that material with comprehension. Ideally, every teacher should have up-to-date test scores showing the ability of each student. With these measures, teachers could match students' ability with the readability of the material. Lacking these conveniences, teachers must take their own measures. One device is the *cloze readability procedure*, an accurate and economical way to test students directly on the materials; we also have some valid and simple *readability formulas* for assessing materials.

Problems with Traditional Direct Testing

The time-honored way to find out if a selection is suitable for students is to test them on it. The teacher identifies some appropriate content material, writes some questions, has the students read the material and try to answer the questions, and then scores the tests. A criterion score (e.g., 75-90 percent correct), used to interpret the students' performances on the test, is supposed to represent a minimum level of desirable performance; the material is considered suitable for those who score at or above the criterion score and unsuitable for those who fall below it.

We want materials to teach content, so theoretically we should try out all materials to see how well they work. However, teachers do not employ direct testing widely and consistently or put much faith in it. Most have had strong scientific, economic, and philosophical reservations about the procedure, which is impeded by problems that make it almost impractical. When teachers do use direct testing, they are nagged by doubts about the results. First, procedures used to write the tests are likely to produce unreliable results. Second, the criterion scores used to interpret the students' scores are based on shaky reasoning and scanty evidence. Third, the tests are so difficult and expensive to make that few teachers could regularly employ the procedure even if they wanted to.

Reliability. Irving Lorge examined the scientific foundations of this procedure and shook teachers' faith in it. In a classic paper on readability (1949), Lorge's argument went something like this:

Anyone who works with test writing knows that it is relatively informal. It is not a science that requires test writers to follow a set of rigid procedures specifying selection and phrasing of items. The test writer may exercise a great deal of freedom in choosing factual content, cognitive processes, and the phrasing of items; all these elements can have a decided effect on test difficulty. Test writers influence the difficulty of tests: two writers making a test over a single passage could produce tests of quite different difficulty, the one writer's test eliciting mostly low scores and the other's mostly high scores. Traditional test writing procedures do not give the test writer enough guidance to avoid this problem.

Teachers often doubt the interpretation of test scores resulting from traditional procedures. Teachers cannot know whether students got high scores because the materials were very effective and taught the students a great deal or whether the test writer merely happened to feel generous that day and wrote an easy test. If the students made low scores, teachers cannot decide with confidence whether the materials were ineffective or the test writer produced a difficult test. Teachers know that the scores on traditionally made tests are influenced by the ability of the student, by the difficulty of the materials, and *by the items that the test writer happened to write for the test.* Teachers are justified in refusing to place much confidence in the results of traditional tests used to evaluate the suitability of instructional materials.

Criterion Scores. The criterion scores used to interpret the results on traditional tests are those recommended by reading authorities (Betts 1946; Bond and Tinker 1967; Harris 1962). These criterion scores differ, depending on the author, but there is some agreement that a student should be able to answer at least 75 percent of the questions on material for supervised instruction and at least 90 percent of the questions on material for voluntary reading. None of the authors offers evidence or describes the reasoning for selecting these particular scores and no one else offers rational support for these criterion scores.

Now, it is certain that each of the authors cited is among the foremost reading experts. Their intuition and common sense may be unexcelled. These criterion scores represent their considered judgments of what constitutes the most desirable performance level for students to reach on materials. The human mind is a marvelous instrument that enabled man to discover many things long before science, but it has also led men to believe that illness was caused by devils that had to be frightened, beaten, or bled out of a person and that thunder was caused by warring gods

hurling each other about. Practice based on common sense and intuition must be examined critically, and verified or refuted scientifically.

As we have discussed, students' scores not only reflect the knowledge they obtained from the material (and therefore the effectiveness of that material for them) but also reflect the idiosyncrasies of the test writer. The size of a score does not provide a reliable estimate of how much of the material's content the student learned. Criterion scores on different tests made from the same material represent different amounts of knowledge and therefore cannot be used to judge the suitability of the material.

Second, criterion scores are based on an assumption of *the-more-the-better* or *up-is-good*. The criterion scores set a minimum level of performance but not a maximum level, thus implying that the better a student performs on a test, the more value he or she is getting from the material. This assumption might be valid if study influences only the amount of cognitive knowledge gained, but the learning situation thus suggested is too narrow to fit the facts: we also expect instruction to influence a student's interests and attitudes. We would not consider material suitable if it were so frustratingly difficult that students avoided it and use of the knowledge it contained, nor would we consider it suitable if students found it so simple, redundant, and boring that they avoided further study of it. Students dislike materials on which their comprehension test scores are either very high or very low and they show a definite preference for materials on which their test performance is at an intermediate level (Bormuth 1971). The criterion scores employed in the traditional procedure set only a minimum score; some of the students who exceed that criterion score on a material may be harmed by its study.

The third reason for questioning these criterion scores is that their authors have not offered any evidence for or explained the reasoning that led them to identify these particular scores. We are left to wonder if the criterion scores sprang from an arbitrary decision that became standard practice simply because of the strength of precedence. The length of the English unit of measure, the foot, took on its value because the king who standardized that measure happened to have a foot of that length. The chief effect of this accident of history was to saddle us with an amazingly inconvenient measurement system.

We use these criterion scores because we believe they will produce more favorable consequences for learning. We believe students will receive more good things from materials if they score

between 75 and 100 percent on them than if they score lower than 75 percent. This may be so, but it is doubtful, since students' willingness to study materials falls off sharply when their scores fall toward the upper end of the 75–100 percent range. We do not know the consequences of relying on these criterion scores, which may interfere with and even subvert the intent of the instruction.

Cost. The traditional procedure for evaluating materials has not received widespread use also because it is costly. The traditional test consumes a sizable amount of teacher time. To produce a test of reasonable quality, the writer must usually take a course in test construction. The traditional test is not an easy thing to write well; the author has found that it requires roughly 8–10 minutes per item to prepare even fairly good questions using traditional test-making procedures. When these questions are tried out and edited to eliminate vague, ambiguous, and mistaken items, the time is at least doubled. To write a 25-item test for a passage, a teacher would probably spend anywhere from 3 to 8 hours. More time is required to type, duplicate, administer, and score the test. Other duties compete for the teacher's time and resources and it should not be surprising that, because of the expense alone, these test-making techniques are not used routinely in more than a few schools.

Traditional procedures for evaluating the readability of materials rely on inadequate and expensive tests and employ criterion scores of dubious quality. Nevertheless, some educators have used these procedures in the past and, I think, justifiably, for they were using the best and only techniques available. The materials chosen for instruction matter, because they differ in effectiveness and thus in the educational benefits students receive from them. Despite the difficulties of evaluating readability, we must improve our techniques wherever possible. And that is what the rest of this paper is about: describing a better way to test the difficulty of instructional materials and evaluate their suitability for students.

Cloze Readability Procedure

A major part of the problem with the traditional procedures arises from the kind of test they use. Those tests are expensive and difficult to make and they produce unreliable measures of the amount a student has learned from the material. We will now consider tests obtained by the cloze readability procedure, a procedure that consists, roughly, of these steps:

First, select one or more passages to represent the materials being evaluated.

Second, starting with the first word in the passage, count the words, marking every fifth one for deletion.

Third, type a stencil of the passage but in place of every fifth word indicate a blank by inserting a 15-space underline.

Fourth, run off copies of this test and administer it without time limits to students. Instruct the students to guess what word was removed to form each blank and to write the word in the blank.

Fifth, score the responses correct only if they exactly match the words deleted.

Finally, figure the percentage of items answered correctly by each student and compare it to the criterion scores that will be presented in a later section.

We will begin by discussing the validity of this type of test and then describe each step in more detail, indicating the research that supports it. A cloze test can be made, administered, and scored in a number of ways, and technically we will be speaking of just one type of cloze test, the *cloze readability test*. We will examine this *cloze readability procedure* in detail and refer to it for the rest of this paper using the terms *cloze*, *cloze procedure*, and *cloze test*.

Validity. When we use cloze readability tests to measure the comprehension difficulties of passages, we assume that these tests measure the mental processes called comprehension and that scores on cloze tests made from different passages measure their relative difficulties. The research is too extensive to detail here, but reviews have been written by Rankin (1965), Bormuth (1967), and Potter (1968).

It seems clear that cloze tests do measure a person's ability to perform the comprehension processes. A cloze test provides students with a passage they have never seen before, but it is a copy of the passage with words systematically deleted. They must use the text that comes before and after each blank to arrive at their answer. Scholars such as Johnson (1966), Miller (1962), and Goodman (1970) studied various aspects of how people process language and consistently found that people use the text they have just read to infer its meaning and structure and to guess the meaning structure of the text about to be read. People also use text just read to confirm their inferences about earlier segments of the text. These processes seem to be integral to the comprehension processes, serving as an essential means of getting information from text as well as furnishing much of the information we get from it. The cloze test taps the comprehension processes at two

points by testing how much knowledge was obtained from the text surrounding the blank and how well the information obtained from the text was employed to obtain additional information. Research shows that we cannot distinguish the processes measured by traditional comprehension test questions from those measured by cloze test items (Bormuth 1969a). These studies show that scores on the two kinds of tests are so highly correlated that the processes that generate those scores are indistinguishable.

The cloze test has important advantages over the traditional comprehension tests. It is made by a relatively simple and mechanical procedure. It is cheaper and easier to make and provides more reliable results. Studies on this subject typically show (Bormuth 1962) that cloze tests of 50 items provide passage rankings with a reliability correlation of roughly .96, or about 92 percent reliable, while traditional comprehension tests have reliability correlations of about .84, or only about 71 percent reliable. Moreover, the difficulty rankings assigned to passages by cloze tests show very high correlations with other measures of passage difficulty, such as those obtained from oral reading word recognition scores, judgments of difficulty by readers, and ease of memorizing. While the correlations reported vary considerably depending on the reliabilities of the various tests used by the various investigators, these correlations are generally about .90 when test unreliability is taken into account. The validity of the cloze test is firmly rooted in basic psychological theory and in the empirical evidence obtained from the extensive research on the subject.

Making a Cloze Readability Test. The rules that distinguish a cloze test from other tests are, first, that the units deleted are words (not letters, syllables, morphemes, or phrases) and, second, that the words deleted are selected by some mechanical process that does not require the test maker to employ ill-defined guidelines such as *"key words"* or *"important concepts."* If teachers want their test results to mean anything, they must adhere exactly to the cloze *readability* test rules which distinguish them from cloze tests in general. We know a great deal about how to interpret the cloze readability tests; any departure from these rules leaves the teacher with uninterpretable results.

Selecting Passages. Though occasionally an entire passage is used, usually we draw samples from the text and assume they give us an unbiased estimate of the difficulty of the entire text. Generally, a text will vary somewhat in difficulty from section to section. Drawing a single passage may give a biased estimate of the over-all

difficulty of the entire test, so it is useful to know just how variable the text is. Up to a point, the more samples tested, the more accurately variability can be determined: any number of passages from 2 to 12 provide the information, but each additional passage beyond 12 yields less and less information. As few as 1 to 3 passages can be adequate to evaluate a text for class use. Selecting a page can be done by noting the total number of pages in the text—i.e., 150 pages—and then randomly selecting a number between 1-150 using a table of random numbers.

Starting with the first paragraph, count down the page 250 words and go to the end of the sentence in which the 250th word appears to get the first sample passage. Carefully inspect the 250-word passage. Look first to see that its comprehension does not depend heavily on the immediately preceding text. If it does, move up one or more paragraphs until this is no longer true and then count 250 words beginning with that paragraph. Dependence, here, means that the passage does not contain anaphoric words and phrases, such as *this, it, many, these words,* or *this criterion,* whose referents are found only immediately preceding the selected part of the text. In some books, everything that follows the first chapter depends on that chapter, so we would select just the first 250 words or so in the text.

Second, the passage should not start in one major section of the text, say in one chapter, and finish in another. Again the solution is to move up paragraph by paragraph in the text until a suitable passage has been obtained.

The text should not contain many mathematical symbols or, for that matter, many numerals. Some scholars are studying the problems involved in using cloze tests on mathematical texts with promising results. But as yet we do not have [illegible] criterion scores for interpreting the results of such tests.

A 250-word passage provides a 50-item test which produces scores that are sufficiently reliable for most of a teacher's purposes. Reliabilities fall close to .85; to increase the reliability by any appreciable amount, say to .90, usually requires a test about twice as long. Cloze tests based on passages 250 to 300 words long happen to fit comfortably on one side of a standard sheet of paper and 50-item test scores convert easily to percentage scores, by multiplying by 2.

Deleting the Words. Randomly choose any one of the first five words in the passage and begin the deletions with that word (unless it is essential to the meaning of the paragraph, in which case begin the deletions with the first inessential word). Mark

every fifth word from that point until 50 words are marked. A "word" is anything set off by blank letter spaces at either end. However, the parts of hyphenated words are deleted as separate words when both parts of the compound word can stand alone as separate words. For example, *self-made* would be counted as two different words since both *self* and *made* can stand alone, but *co-operate* would be counted as a single word since the prefix co- cannot stand alone. A number such as *1584* is deleted as if it were a single word. Punctuation and hyphens are never deleted, but apostrophes are deleted along with the word they appear in.

The decision to delete every fifth word is based on both scientific and practical considerations. Leaving only one or two words between blanks in a cloze test reduces test efficiency. The answer to one item often depends upon whether the surrounding items were answered correctly. Each item measures the same thing that the surrounding items measure and yields little new information. MacGinitie (1960) showed that this dependency among items drops about as far as it can drop when the distance between cloze blanks is increased to four words of text. Deleting every fifth word gets us close to the maximum efficiency per item and permits using a short, economical passage that produces reliable test scores.

Reproducing the Test. Set the typewriter margins for ordinary manuscript typing: a 1½" left margin and a 1" right margin. Set the line ratchet for either a space and a half or two spaces between lines. Indent paragraphs and set other margins in the manner dictated by ordinary manuscript format. Type the passage onto a duplicating stencil. Omit each word marked for deletion and type a 15-space underline instead, leaving spaces before and after the blank just as you would for a word. Leave no space between a blank and hyphens or between a blank and punctuation that follows it.

Underlined blanks must be of a standard length because variations in the underlining affect the scores. Scores on cloze readability tests are meaningless to a teacher unless they can be compared with the criterion scores that will be presented later. These criterion scores were shown by rational procedures to represent the most desirable levels of performance on cloze tests of this sort. But these criterion scores were found for tests made only by the procedure being described here; if anything is done to a cloze test to make it either easier or harder than tests made by this standard procedure, then neither these criterion scores nor any others are valid for interpreting scores on the test. If the blanks are varied to correspond to the length of the word, if the number of

underline strokes tells how many letters are in the word, or if the response is cued by any means other than the standard underlined blank, then the test is invalid for evaluating readability.

Fig. 1. Sample cloze test with instructions to students

Instructions

At the bottom of this page is a sample of a new kind of test. Each of these tests is made by copying a few paragraphs from a book. Every fifth word was left out of the paragraphs, and blank spaces were put where the words were taken out.

Your job will be to guess what word was left out of each space and to write that word in that space.

It will help you in taking the test if you remember these things:

1. Write only one word in each blank.
2. Try to fill every blank. *Don't be afraid to guess.*
3. You may skip hard blanks and come back to them when you have finished.
4. Wrong spelling will not count against you if we can tell what word you meant.
5. Most of the blanks can be answered with ordinary words but a few will be
   - numbers like. . . . . *3,427* or *$12* or *1954*
   - contractions like. . . . . *can't* or *weren't*
   - abbreviations like. . . . . *Mrs.* or *U.S.A.*
   - parts of hyphenated words like. . . . . *self-* in the word *self*-made

Sample Test

Below is a sample of one of these tests. Fill each blank with the word you think was taken out. You may check your paper when you finish it by looking at the answers which are written upside down at the bottom of the page. Write neatly.

The Beaver

Indians call beavers the "little men of the woods." But they _____ really so very little. _____ beavers grow to be _____ or four feet long _____ weigh from 30 to _____ pounds. These "little men _____ the woods" are busy _____ of the time. That _____ why we sometimes say, "_____ busy as a beaver."

_____ know how to build _____ that can hold water. _____ use their two front _____ to do some of _____ work. Cutting down a _____ with their four sharp-_____ teeth is easy. A _____ can cut down a _____ four inches thick in _____ 15 minutes.

Answers: 1. aren't 2. Most 3. three 4. and 5. 40 6. of 7. most 8. is 9. as 10. Beavers 11. dams 12. They 13. paws 14. their 15. tree 16. pointed 17. beaver 18. tree 19. about

Test Instructions. Test instructions worked out by the author (Bormuth 1964) are accepted as standard for administering these tests. The instructions are shown in figure 1. They are read to the students as the students read them silently. After students have several experiences with these tests, it is usually unnecessary to repeat instructions to them in subsequent testings. Warn students before they take their first cloze test that it will be harder than the other tests; this reduces the effect of students' anxieties on their scores. Formal time limits are not imposed on the tests. The students are encouraged to work as long as they think they can improve their scores, but some children work past the point where their efforts are productive. The tester can stop testing when that student's score appears to be a valid measure of his or her ability. Variations from these procedures reduce the validity of the test scores for evaluating readability.

Selecting Students. When evaluating materials for use in one classroom, administer the tests to all students to determine which students should or should not use the text. When evaluating materials for system-wide use, administer the tests to students who represent the full range of social, economic, and ethnic groups in the community. Dividing the school population into these categories, randomly draw a sample of students from each category. Seek the aid of a research technician in designing this sampling procedure.

Typically, children below grade four still struggle with the basic word recognition skills and test scores not only reflect understanding but also how well they are able to pronounce and recognize words. Curriculum experts often think a text should be used primarily to exercise basic reading skills and only secondarily to provide subject matter information. To evaluate these texts, test the student's ability to recognize the words rather than ability to comprehend.

Scoring the Test. A student's responses are scored correct only when they exactly match the words deleted. Minor spelling errors are scored correct so long as the response is otherwise correct. However, omission of plural or tense endings is scored incorrect, as in *table* for *tables* or *work* for *worked*. Convert raw scores to percentage scores by multiplying the number correct by two, when a 50-item test is used, or by dividing the number correct by the total number of items in the test and multiplying by 100, for tests of other lengths. To save time, use a window key with holes cut in a sheet of paper so that everything but the responses is masked

when the sheet is placed over the test. Write the correct response beside each window.

Some argue that when a student writes in a synonym for the deleted word this shows that he or she understood the passage and therefore synonyms should be scored as correct. Their speculations are correct in asserting that a synonym reflects comprehension. Scores based on synonyms show a fairly high correlation with other measures of comprehension, such as scores on standardized tests of comprehension achievement and cloze test scores based on exactly matching responses. The chief problem is that scorers disagree on what is an acceptable synonym and, so far, it has been impossible to devise a set of rules for settling these disputes. Consequently, scoring synonyms correct detracts from the reliability of the scores. Studies repeatedly show that scoring synonyms correct adds a few points to everyone's score but does not increase the validity of the tests. When synonyms are scored correct, it takes from two to four times as long to score the tests, making the testing operation much more expensive.

Scholars settled on the practice of scoring only the exactly matching responses correct. The criterion scores we will examine next are based only on tests scored in this manner. *Test scores obtained in any other way are invalid for use with those criterion scores, and therefore, for use in evaluating the readability of texts.*

Cloze Criterion Scores

Suppose we have cloze test scores for several books we are considering for adoption. We must assign values to the scores saying how *good* or how *valuable* each score is. How should we assign values to scores on tests of this type?

It is customary on standardized tests to assume that knowledge is good, that more knowledge is better, that children are in a race to get knowledge, that test scores reflect how well the children are running the race, and that the value of a score depends on how it compares with the scores of other students. These tests are accompanied by percentile norms, standard score norms, and grade-placement norms which do not tell us either the quantity or usefulness of the knowledge to the student in performing tasks such as reading a newspaper or deciding which of two cans of beans represents the best buy. These norms merely tell us how far he or she is running ahead of or behind the pack.

This method of assigning values to test scores is not the only way values can be assigned and it is not very useful for assigning values to scores on cloze readability tests. A high score is not

always better than a lower score. The easiest passages produce the highest cloze scores and the highest comprehension test scores. Part of the reason students get high comprehension scores is that much of the content of the passages was learned previously. A higher cloze score may not be a *better* cloze score when we consider only the information a person gains from a passage.

Children do not like very difficult materials because they are frustrating and unpleasant to try to learn from. This would suggest that higher cloze scores are better, except that students also dislike materials that are very easy for them. Neither very high nor very low cloze readability scores are desirable, but which score between 0 and 100 percent *is* the most desirable one?

Rationale for Selecting a Criterion Score. Selection of the *most desirable* score requires a definition of desirable. The value of a score depends on what that level of performance gets us and how much we value the things it gets for us. When we read a text, it is because we want to get something in return for our efforts: perhaps aesthetic enjoyment, perhaps information, or perhaps no more than a painless way to use up some portion of our lives. Here we are interested in: (1) how much information the average person gains from the text; (2) how willing he or she is to read it; (3) how novel the content is to the person; and (4) how rapidly he or she is able to read it. These four measures of the suitability of a text determine the value of each level of performance on that text. (See Bormuth 1971 for a complete account of the studies underlying this reasoning.)

Can we determine the value of our cloze score simply by finding how much of each of the four attributes we get when we obtain a score of a given size? Is the value of this score just the sum of these amounts? and is the best score the score at which this total is the greatest? Almost—but not quite. We cannot just add these things together without taking into account how much people value them. We must find the *relative values* that people place on each of the four outcomes of reading a text and add them to determine the cloze score with the greatest value. This will be our criterion score.

Our conclusions are limited because people read for a variety of reasons and place different values on what they get from reading in various situations. No single criterion score can serve all purposes; a score serves only the purposes for which it was designed and is useful to people who share similar values. Although not all people value information gain and the other benefits of reading in the same way, people fortunately are remarkably

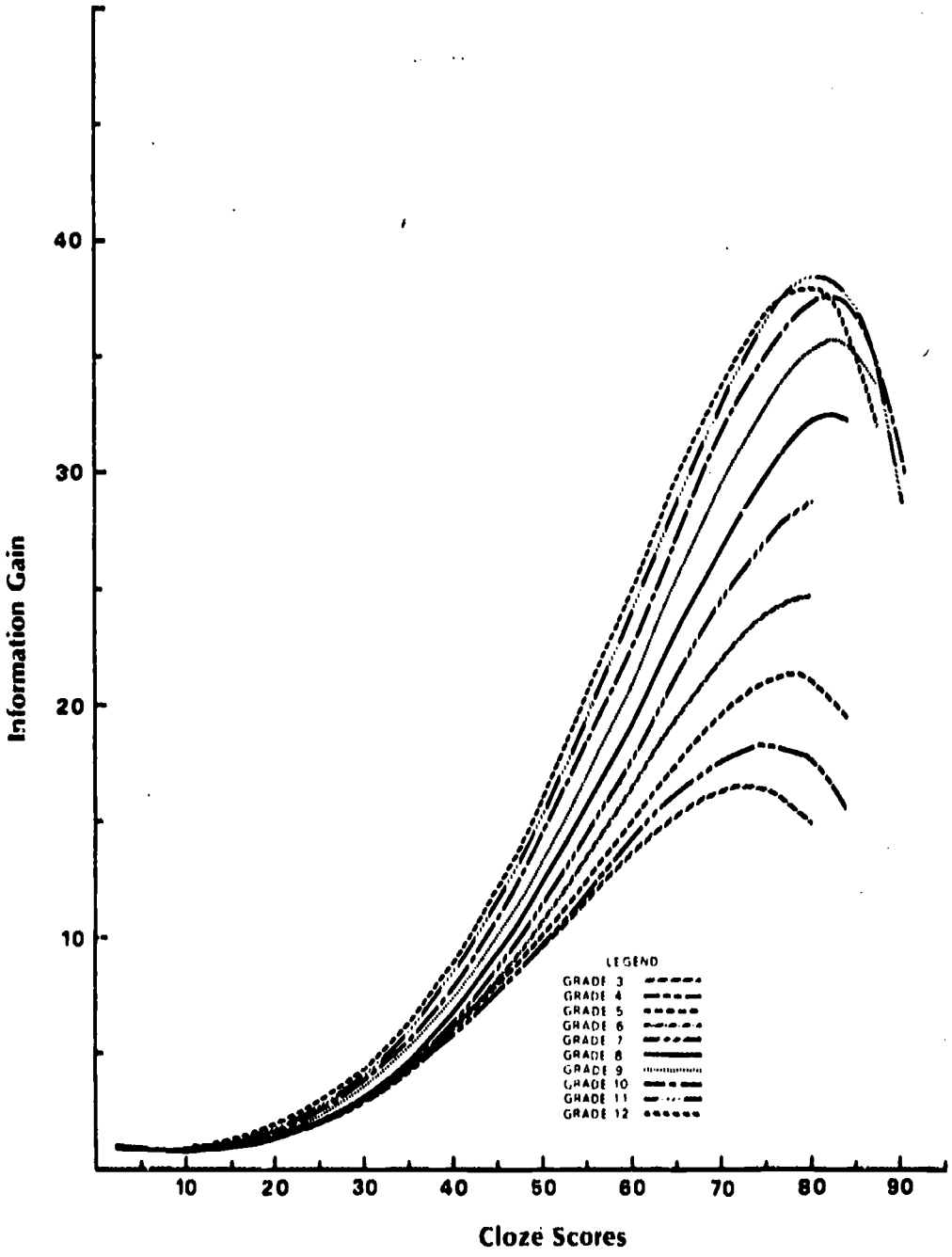


Fig. 2. Information gain measured by performance on completion tests

similar, and the criterion scores to be given here represent the consensus of the vast majority of the teachers involved in these studies.

To determine criterion scores for cloze readability tests then, we need to find the relationships between cloze scores and what we want to get from reading. The four outcomes mentioned earlier were studied: information gain, a student's willingness to study a material, novelty of the material's content, and rate of reading.

Information Gain. Information gain was measured using several passages that were nearly identical in difficulty. A traditional comprehension test was made for each passage. Students were first given one of the tests and told to guess as many of the answers as they could; they were not given the passages on which the tests were based. The purpose was to find out how much students already knew about passages at that level of difficulty. Each student was then given a passage of matched difficulty to read and took the comprehension test made for it. Each student's score on the guessing test was subtracted from his or her score on the actual comprehension test to obtain a residual gain score. Finally, students took a cloze test over a third passage that was matched for difficulty with the other two passages. The passages and tests were switched for each student to assure that peculiarities of a particular passage could not bias the results. Eight sets of passages were used so that eight difficulty levels and a variety of subject matters were represented. A total of 1600 students were tested, 160 at each of the grade levels 3 through 12.

Figure 2 shows a graph of the results of the information gain tests. These lines show the average amount of information the students gained at each level of cloze performance. The horizontal scale shows the students' cloze scores and the vertical scale shows the amount of information gain. Each line shows the results for students in a different grade level. Students who made cloze scores as low as 0 to about 35 percent gained little or no information from the passages. These students had much difficulty understanding the passages. Thereafter, students who made increasingly higher scores made increasing gains in information. But this rise halts when we reach students who got cloze scores of about 80 percent. After that point there is a sharp drop-off in the amount of information students were able to gain. A cloze criterion score based on information gained from a passage would be set at 80 percent, the point where students receive the most of this value—information—from their efforts.

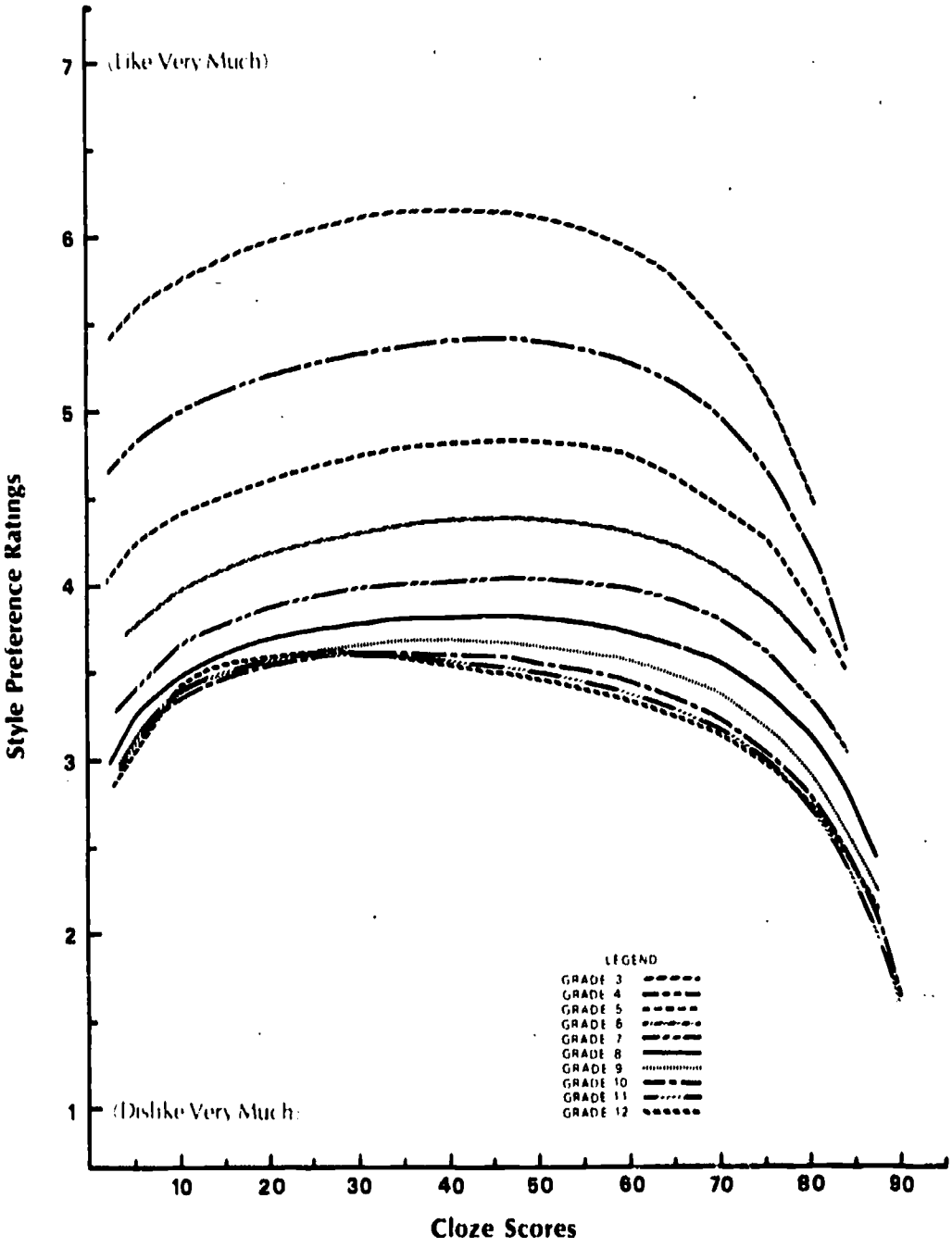


Fig. 3. Willingness to study measured by style preference ratings (curve for textbook reading)

However, the matter is not that simple. When the data are analyzed for students at each grade level separately, the curves are parallel but different, as figure 2 shows. The graphs of the other tests also show that the curves are different for each grade level. This tells us that we must use different criterion scores for students at each grade level.

Willingness to Study. Cloze scores were measured on one passage and willingness to study on a passage of matched difficulty. Willingness to study was measured using a seven-point scale with 7 described as "like very much" and 1 described as "dislike very much." The student read the passage and rated it on this scale. Figure 3 shows the results. Students rejected both very difficult and very easy study materials. They preferred passages on which their scores fell into an intermediate range of difficulty. Their preferences reached a peak well below a cloze score of 80 percent where information gain is at a maximum. Students were also asked to rate the passage for use as a textbook, a reference book, and as voluntary reading. These curves were similar in overall shape, but peak cloze scores differed. We must use different criterion scores for students at each grade level as well as for each major use to which a text is put.

Difficulty. Students also rated these passages on a difficulty preference scale which seemed to measure how new or novel the content of the passage was to them. One end of this scale was described as "not suitable" and the other end as "suitable." Figure 4 shows this curve when the students were rating the passage as a textbook; 2.0 represents "not suitable" and 4.0 represents "suitable." The curves for reference reading and voluntary reading roughly paralleled this one but differed at the peak points.

Rate of Reading. Students were timed as they read the passage on which they took their comprehension test and these times were converted to words-per-minute scores. Figure 5 shows the results when this curve is plotted.

Weighting the Reading Outcomes. Teachers furnished judgments on the relative value of each of the four outcomes. Given a description of each outcome and a 10-point scale, teachers chose the least valued outcome and assigned it a "1" rating, assigned the most valued outcome a "10" rating, and placed the remaining outcomes at points that seemed most appropriate to them. A total of 101 teachers, drawn in about equal numbers from each grade level, made the ratings. The grade level at which a teacher worked

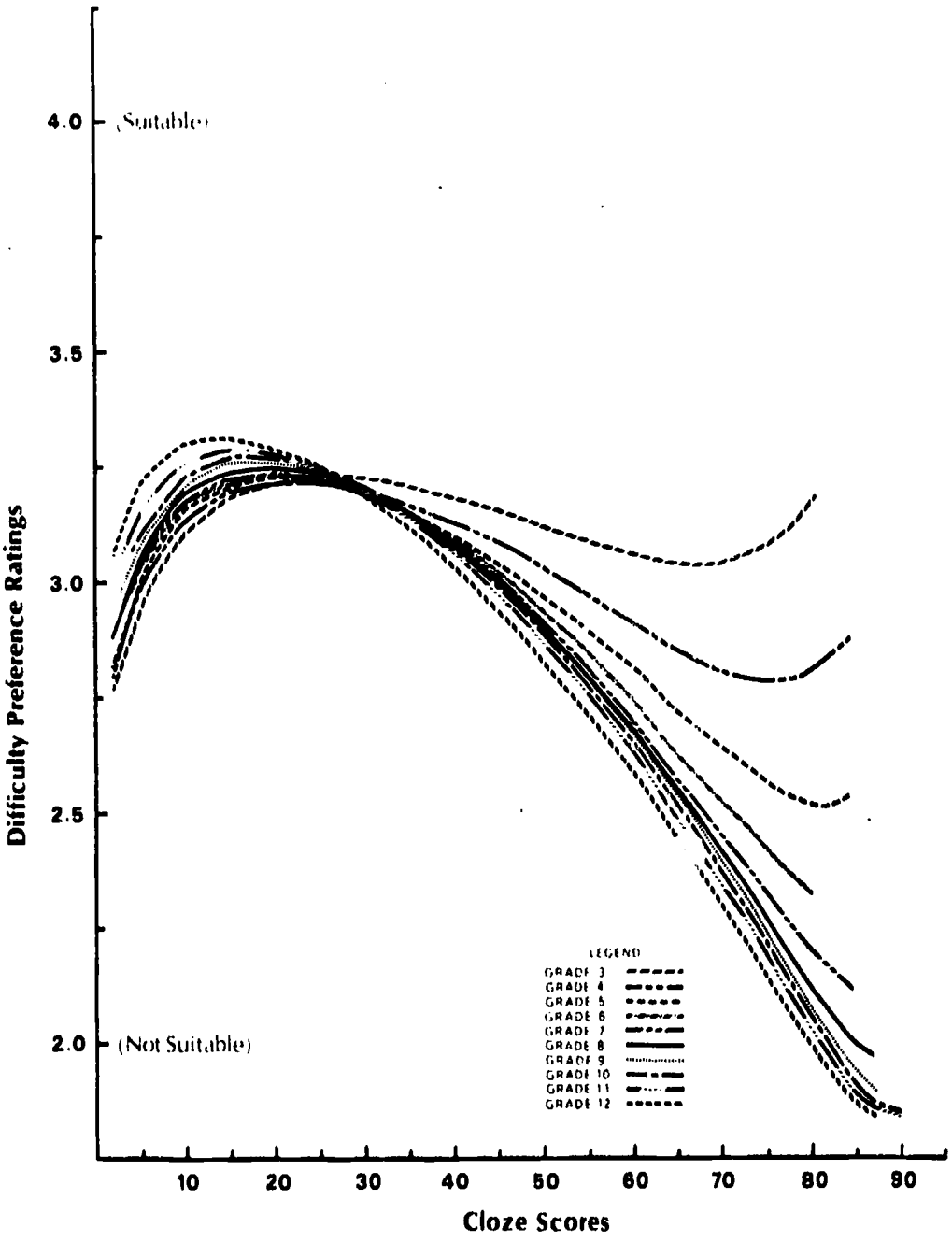


Fig. 4. Difficulty measured by difficulty preference ratings (curve for textbook reading)

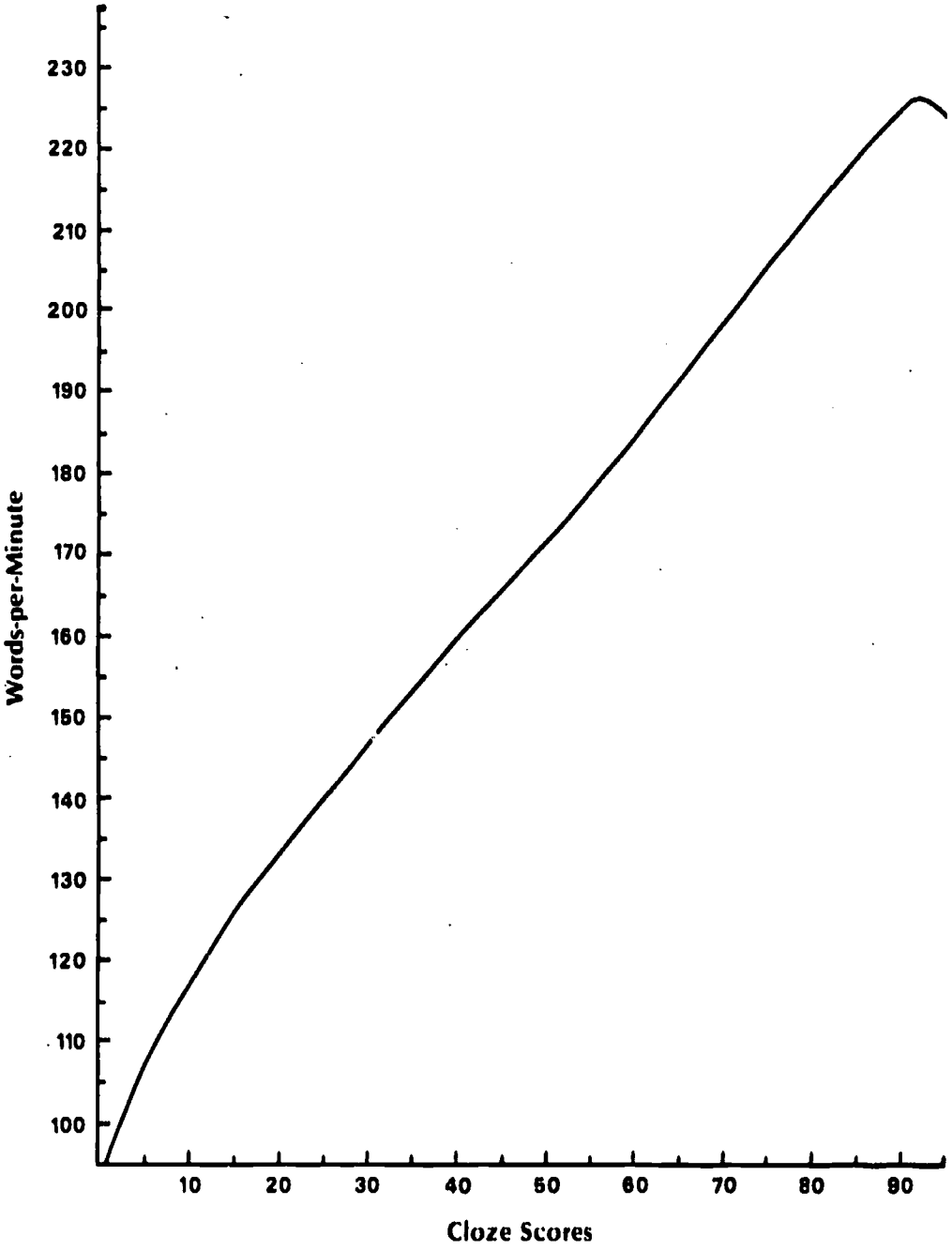


Fig. 5. Rate of reading measured in words-per-minute (combined curve for grades 3-12)

had no significant effect on the ratings, but the use to be made of the materials had a great effect. "Information gain" earned a mean rating of 5.5 when teachers were rating its importance for textbook reading, but earned only 2.7 when they rated it for voluntary reading.

Finding the Criterion Scores. The last step consists of using the outcome ratings to calculate the value of each cloze score and then identifying the most valued score as the criterion score. All the scales were converted to standard scores so they would be expressed in roughly identical counting systems. The information gain that correlated with a cloze score of one percent was identified and multiplied by its rating for that type of reading task (textbook, reference, or voluntary); this weighted sum indicates the information gain of a cloze score of one percent. The values that this same cloze score obtains from our other measures of suitability—willingness to study, difficulty, and rate of reading—were calculated in the same way. These four values were then added together to obtain the total value of a cloze score of one percent.

This process was repeated for a cloze score of two percent, three percent, and so on until the total value of all cloze scores had been found. The score receiving the greatest total value was taken as the criterion score. Thirty criterion scores were established, one for each of the three uses of the material within each of ten grade levels (see table 1). I hesitated to include the results for grade three because I have rather strong reservations about whether the materials at that grade level should be evaluated this way. However, they are included along with the admonition that,

Table 1

Criterion Scores for Cloze Readability Tests

Use to Be Made of the Materials	Grade Level of the Student									
	3	4	5	6	7	8	9	10	11	12
Textbook	59	58	57	56	55	53	52	50	50	49
Reference	55	53	52	51	49	48	47	45	45	44
Voluntary	90	62	54	50	49	46	44	40	34	34

for children below grade four, it is probably as important, or even more important, to evaluate the word recognition difficulty of the materials as it is to evaluate their comprehensibility.

These criterion scores are far superior to any other criterion scores available. Those previously offered were selected for unknown reasons and had unknown consequences; the ones shown in table 1 are based on a logical model that explicitly incorporates our social values and combines them with scientific evidence. As the results of the first effort to establish rational criterion scores, they are only approximations of the criterion scores we will eventually identify, although future research will probably alter them only slightly. We cannot accept any criterion score unless we can accept the reasoning and evidence on which it is based. Only the barest sketches of reasoning and evidence have been given here, and I strongly recommend that users of these criterion scores study the original research report (Bormuth 1971) carefully and critically to assure themselves.

Evaluating Cloze Readability Scores. How do we determine if material is suitable for a student and how do we determine if it is suitable for a group of students?

With a single student who has taken one cloze test over each of several sets of materials, we set the highest value on the material on which his or her test score most closely approaches the criterion score and correspondingly lower values on the other materials tested. Presumably, we have also examined and evaluated the contents of these materials for quality on some scale. The material chosen should best represent a compromise between content and readability. Readability and content values have lower limits of acceptability. In readability, we might place this lower limit at a cloze score of 35 percent. Figure 1 shows that students scoring below 35 percent exhibit less than a five percent information gain. Hence we might say that materials falling below that level are essentially worthless for teaching that student and probably should not even be considered in the final selection.

With several tests over a text eliciting two or more scores, we can estimate roughly how variable its readability is. Some variability is unavoidable and possibly desirable for breaking monotony. Too much variability almost assures that part of the material will be poorly learned by the students and the teacher will have to find the difficult spots and supplement the text with explanations. A text of even readability is preferable over a text of quite variable difficulty.

Ideally, in selecting texts for a large number of students who vary in reading ability, we would not choose a single text because great differences among individuals make any single text poorly suited to a large number of students. But texts differ in content and organization and teachers lack the time to vary instruction accordingly. It is often necessary to employ one text with students who have a wide variety of abilities. To select such a text, choose material on which the mean cloze score falls somewhat above the criterion score, assuring that the material is partially comprehensible to nearly everyone. Some of the able students will find this text insipid and repetitious of previous material. But these students can usually work independently, reading more advanced references and doing other projects. An easy text helps the slower students extract the basic content, relieving the teacher of the obligation to repeat the content through lectures. It also provides slower students with practice to extend their reading skills.

Formulas for Predicting Readability

If we want to know how well our students will do on a material, we merely need to make a set of cloze readability tests to find out. The cloze readability procedure is direct and it is buttressed by research that is possibly the most thoroughly validated and sophisticated method of testing presently used in education. However, though cheaper than the procedures it replaces, it is still expensive. It still requires a lot of work to make, administer, score, and interpret these tests—often, more work than the teacher can justifiably devote. Consequently, the teacher needs an even less expensive way to determine the difficulty of materials. The readability formula is such a device.

Scientists studying the sources of difficulty found that difficulty is regulated by the language of a passage. They published formulas that permit teachers to predict the difficulty of materials by measuring a couple of features of the language in the materials, inserting these numbers into the formula, and then working out the formula. These formulas have steadily improved in accuracy and in the ease with which teachers can use them; we will present one of the most accurate and simple.

Derivation of the Formula. A complete description of the method and the data on which the formula is based can be found in Bormuth (1969b). The passages selected should be representative of the kind of passages we want to evaluate. From instructional materials, 330 passages were selected following much the same rules given when we discussed making cloze tests, except that no

two of these passages came from the same text. We chose 33 passages from each of 10 subject matter areas with materials from grade one through college.

Cloze difficulties were determined and converted to grade-placement scores, using cloze readability tests made from each passage. These were administered to groups of 285 students from grade 4 through grade 12; each grade level was represented about evenly in a group. No single child took 330 tests, but 50 carefully matched groups of tests were used.

Various features of the language in the passages were analyzed and measured. Features used to predict readability range from counts of the average number of letters in the words or sentences in a passage to counts of the average number of syntactic transformations required to go from the surface structure of a sentence to its deep structure kernels. The variables included in a readability formula reflect how simple or foolproof we want that formula to be. It is possible to obtain a very accurate formula with a validity as high as .93; it is nearly impossible for even a clever writer to fool such a formula by making a passage appear easy when, in fact, it is difficult. But to use such a formula we must pay a price. The price is the cost of applying the formula, for very foolproof and accurate formulas include a large number of variables that are often complex and time-consuming to analyze. However, we can use a simpler formula if we sacrifice a small amount of accuracy and assume the writer did not try to make the passage appear easier than it really is.

The two variables selected, word length and sentence length, are easy to analyze without getting involved in word lists or complex grammar analyses. We chose to measure length in terms of the number of letters a word or sentence contains, avoiding decisions about how many syllables a word contains. These two variables proved to be as accurate as any other variables used so far to predict readability.

Calculating the formula involved a mathematical procedure called polynomial multiple-regression analysis, but the result is a fairly simple equation that we all learned to solve in first-year algebra:

$$d = 1.069 - \left(.106\frac{l}{w}\right) - \left(.0036\frac{l}{s}\right) + \left[.000002\left(\frac{l}{s}\right)^2\right]$$

This is the basic readability equation that we will use. However, table 2 is provided to eliminate the need to solve either this equation or its modified form. What this equation says is: The

difficulty, *d*, of a passage in terms of cloze means is equal to *1.069* minus *.106* times the average number of letters per word in the passage, *l/w*, and so on. The symbol *l/s* stands for letters per sentence.

This basic formula can be used to obtain modified formulas that express passage difficulty in terms of grade placement scores: *the grade placement score a student would have to get on a certain standardized achievement test in order to reach a criterion score on that material.* This group of formulas has a validity of about .81. Two-thirds of the time a passage of about 100 words predicts a student's cloze score to within about 6 cloze percentage points, or about .8 of a year in grade placement terms. Two passages drawn from the same material increase the validity to about .89 and five such passages increase validity to about .94.

Applying the Formula. Table 2 was prepared to avoid the need for calculation using the formula. The top third of table 2 is consulted when teachers want to use 35 percent on a cloze test as their criterion score, the middle third consulted for a criterion score of 45 percent, and the bottom third for 55 percent. Each number in the body of this table stands for a grade placement score on a standardized test of reading achievement; the number 6.3 in the upper left hand corner, for example, stands for the third month of the sixth grade. The numbers down the left hand column of the table stand for the average number of letters in the words of a passage and the numbers across the top of the table stand for the average number of letters in the sentences.

The readability of a passage can be determined by calculating the average number of letters contained in the words and sentences of the passage. Choose a passage in the text you are considering according to the rules given when we discussed how to make a cloze readability test. Count (1) the number of letters in the passage, (2) the number of words, and (3) the number of sentences. Divide the number of letters by the number of words to determine the average number of letters per word in your passage. In table 2, in the section labeled with the criterion score you've decided to use, find the "letters per word" (left-hand column) closest to the average letters per word of your passage. Then divide the number of letters by the number of sentences to determine the average number of letters per sentence and find the number closest to your average in the "letters per sentence" heading at the top of the table. The number at the intersection of your "letters per word" row and "letters per sentence" column is

Table 2

Difficulties of Passages in Terms of Grade Placement Scores for Various Combinations of Word Length and Sentence Length

Letters per Word	Letters per Sentence										
	20	42	64	86	108	130	152	174	196	218	240
Criterion Score = 35%											
3.4	6.3	6.8	7.5	8.0	8.3	8.6	8.8	8.6	8.3	7.9	7.5
3.6	6.6	7.1	7.8	8.3	8.7	9.1	9.1	8.9	8.6	8.3	7.8
3.9	6.9	7.5	8.2	8.8	9.1	9.4	9.4	9.3	9.1	8.7	8.1
4.1	7.2	7.8	8.5	9.1	9.4	9.7	9.7	9.7	9.4	9.1	8.5
4.4	7.5	8.2	8.9	9.5	9.7	10.0	10.0	10.0	9.7	9.4	8.8
4.6	7.8	8.5	9.2	9.7	10.0	10.3	10.3	10.3	10.0	9.7	9.2
4.8	8.1	8.9	9.5	10.1	10.3	10.6	10.6	10.6	10.3	10.0	9.5
5.1	8.4	9.2	9.8	10.3	10.6	10.9	10.9	10.9	10.6	10.3	9.8
5.3	8.8	9.6	10.2	10.6	10.9	11.2	11.3	11.1	10.9	10.6	10.1
5.6	9.2	9.9	10.5	10.9	11.2	11.5	11.5	11.3	11.2	10.8	10.4
5.8	9.5	10.2	10.8	11.1	11.5	11.6	11.6	11.5	11.4	11.1	10.7
Criterion Score = 45%											
3.4	7.8	8.4	9.0	9.3	9.7	9.8	9.9	9.8	9.6	9.2	8.8
3.6	8.1	8.7	9.2	9.6	10.0	10.1	10.2	10.1	9.9	9.5	9.1
3.9	8.3	9.0	9.5	9.9	10.3	10.4	10.5	10.4	10.2	9.8	9.4
4.1	8.6	9.2	9.7	10.2	10.6	10.7	10.8	10.7	10.5	10.1	9.7
4.4	8.8	9.5	10.0	10.5	10.9	11.0	11.1	11.0	10.8	10.4	10.0
4.6	9.1	9.7	10.3	10.8	11.2	11.3	11.4	11.3	11.1	10.7	10.3
4.8	9.3	10.0	10.6	11.1	11.5	11.6	11.7	11.6	11.4	11.0	10.6
5.1	9.6	10.3	10.9	11.4	11.8	11.9	12.0	11.9	11.7	11.3	10.9
5.3	9.9	10.6	11.2	11.7	12.1	12.2	12.3	12.2	12.0	11.6	11.2
5.6	10.1	10.9	11.5	12.0	12.4	12.5	12.7	12.5	12.3	11.9	11.5
5.8	10.4	11.2	11.8	12.3	12.8	12.8	13.0	12.8	12.7	12.2	11.8
Criterion Score = 55%											
3.4	9.3	9.8	10.2	10.6	10.8	10.9	11.0	10.9	10.8	10.5	10.1
3.6	9.5	10.0	10.4	10.8	11.1	11.2	11.3	11.2	11.0	10.8	10.4
3.9	9.7	10.2	10.7	11.1	11.4	11.5	11.6	11.5	11.3	11.0	10.6
4.1	9.9	10.4	10.9	11.3	11.7	11.8	11.9	11.8	11.6	11.3	10.9
4.4	10.1	10.7	11.2	11.6	12.0	12.1	12.2	12.1	11.8	11.5	11.1
4.6	10.3	10.9	11.4	11.8	12.3	12.4	12.5	12.4	12.1	11.8	11.4
4.8	10.5	11.2	11.7	12.1	12.6	12.7	12.8	12.7	12.4	12.1	11.6
5.1	10.8	11.4	11.9	12.4	12.9	13.0	13.1	13.0	12.7	12.4	11.9
5.3	11.1	11.7	12.3	12.7	13.2	13.3	13.4	13.3	13.0	12.7	12.2
5.6	11.3	11.9	12.6	13.0	13.5	13.6	13.7	13.7	13.3	13.0	12.5
5.8	11.6	12.3	12.9	13.3	13.8	14.0	14.0	14.0	13.7	13.3	12.8

the readability of your passage expressed in a grade placement score. Repeat this process on as many passages from the text as necessary to get the accuracy you want and can afford to take the time to get.

Now let's go through this procedure again using the first paragraph from this paper as our example. Following the rules discussed earlier, imagine that we have drawn a sample that happens to coincide with a paragraph (in actual practice, we can ignore paragraph boundaries). Let's say that we want to know how much reading ability is required to perform on this passage at both the 35 percent and the 55 percent criterion levels on a cloze text.

By counting we find that there are 622 letters, 120 words, and 5 sentences in the sample paragraph. When we divide the 622 letters by the 120 words, we find that the average length of a word is 5.18 letters, so we mark the 5.1 rows in table 2 for the two criterion scores we are interested in.

When we divide the 622 letters by the 5 sentences, we find that the average sentence length is about 124 letters, so we mark the column headed 130, the closest column to this average sentence length. When we go down this column in the 35 percent criterion section to find where it intersects our word length row, we find the number 10.9. When we do the same for the 55 percent criterion, we find 13.0. For greater accuracy, we might want to repeat this process for several other passages *from* the paper.

Using this table rather than the formula reduces the amount of calculation required. The major source of labor, counting letters, can be reduced further by estimating, since adding or subtracting a few letters would not change the results. To estimate, count the number of letters in each line in the first two sample passages to be analyzed; omit partial lines and indented lines. Add these counts and divide by the number of lines to get an estimate of the average number of letters per line. On succeeding passages, estimate the number of letters by first counting the number of full lines and multiplying by the average number of letters per line, and then adding to this the number of letters actually counted on the partial lines.

Interpreting the Results. If a teacher took the results at face value, he or she would go to the test records for the class and look up each student's score on a recent reading achievement test. The teacher would not give this paper to any student with a score of less than 10.9 on the reading achievement test. Table 2 shows that those students are likely to gain little or none of the information the paper contains.

Similarly, students with scores between 10.9 and 13.0 should be introduced to the paper before they are required to read it. The teacher should go through the paper carefully, noting the vocabulary items that will cause difficulty and making an outline of the main ideas. The teacher should teach these concepts, discuss the major points of the paper, and possibly even present the students with questions to be answered when they finish. Although most students would benefit from this kind of introduction, it is essential for students at these reading levels. Students with reading achievement scores at or above 13.0 can probably read well enough to get most of the content without help.

When several months have passed since a student took the standardized reading achievement test, the score will underestimate his or her present ability level. To correct for this quickly, determine how many school months have passed since the student took the test, count each month as .1 of a year, and add this amount to the score. A more accurate procedure consists of subtracting one year from the student's grade placement score on the test and dividing this number by the number of years and months spent in school up to the time the test was given (months in school are scored as .1 of a year). This number estimates the rate of past learning of reading skills in skill years per school year. Multiply this rate by the number of years and months of school up to the present time; then add one year to arrive at the student's estimated present grade placement score. (One year is subtracted from a student's grade number to get the number of years in school, assuming that the student was neither retained nor accelerated.) This estimate of the student's present score can be compared with the passage readability score. The teacher should be alert that a few students will always show scores somewhat higher or lower than their true ability warrants.

The reliability of the formula can be increased by analyzing more than one sample passage. As the number of samples from a text reaches six, however, the reliability and validity of the estimate rise extremely high (both in excess of .94). Little can be gained by analyzing more passages.

The teacher should carefully note variations in readability from one sample passage to another. Most variations will lie in a range of about one year. When they spread out more, select a more uniform text or locate the difficult spots and give the students special preparations before assigning those sections.

The original study on which this formula is based indicates that grade placement numbers derived from scores from the *California*

Reading Test, 1963 norms, are inflated grade placement scores compared to scores from other tests. The inflation amounts to about one year. If a school uses the *California Reading Test*, 1963 norms, the readability scores in table 2 accurately predict a student's performance on a text. But if a school uses another test, subtract about a year from the passage readability score determined by table 2. The United States Office of Education is conducting an elaborate study to compare the norms of all the major reading achievement tests; when it is complete, these passage readability scores should be freely convertible to any test.

Summation

Written materials are one of our most important instructional devices. If students can understand them, they relieve the teacher of the laborious and time-consuming task of delivering content through lectures. This frees the teacher to spend time directing the students in discussions and other activities that help them evaluate their knowledge, apply it to problems of current interest, and even learn how new knowledge is developed. If the instructional materials are incomprehensible, the teacher's main concern becomes one of staving off academic disaster. This means long hours spent preparing and delivering lectures and explanations of the basic content. The teacher cannot lead lively discussions and projects on content that the students have not yet learned, and knowledge of basic content is normally considered the minimum essential of education. Large amounts of educational funds have been wasted on materials that could not perform their role. We must assure that students are literate on the materials they must study.

One approach is to teach the students as many reading skills as may be required by their instructional materials. The second approach is the concern of this paper, to alter the materials so that they more closely match the reading skills of the students. Teachers cannot be expected to rewrite materials, but they can select materials. This paper describes two methods for evaluating the difficulty of materials. The first method consists of making cloze readability tests over materials and testing the student on them directly. The second consists of applying a readability formula to the materials. The cloze readability testing procedure is more direct, but considerable amounts of teacher and student time are required, even though it is more economical than earlier methods of assessing texts directly.

A readability formula, on the other hand, has problems that the cloze readability testing procedure does not, although it can predict cloze readability scores with sufficient precision for teachers more economically than can cloze readability testing. However, the formula is precise only under certain conditions: when the students recently took a standardized reading test, when that test was the same one used to derive the readability formula, and when the teacher has analyzed several samples from the text. These conditions are hard to match exactly, but teachers can compensate for them as we discussed. In any case, teachers have here two devices for increasing the effectiveness of the materials chosen for their students.

REFERENCES

Betts, E. A. *Foundations of Reading Instruction*. New York: American Book, 1946.

Bond, G. L., and M. A. Tinker. *Reading Difficulties: Their Diagnosis and Correction*. New York: Appleton-Century-Crofts, 1967.

Bormuth, J. R. *Cloze Tests as a Measure of Comprehension Ability and Readability*. Ph.D. dissertation, Indiana University, 1962. Ann Arbor, Michigan: University Microfilms, No. 63,2586.

———. *Relations between Selected Language Variables and Comprehension Ability and Difficulty*. USOE Cooperative Research Project No. 2082. Los Angeles: University of California, 1964.

———. *Implications and Use of Cloze Procedure in the Evaluation of Instructional Programs*. Los Angeles: University of California Center for the Study of Evaluation of Instructional Programs. Occasional Report No. 3, 1967.

———. "Factor Validity of Cloze Tests as Measures of Reading Comprehension Ability." *Reading Research Quarterly* 4/3 (1969a): 358-65.

———. *Development of Readability Analyses*. USOE Technical Report No. 7-0052. U. S. Office of Education, 1969b.

———. *Development of Standards of Readability: Toward a Rational Criterion of Passage Performance*. USOE Technical Report No. 9-0237. U.S. Office of Education, 1971. (Also available through ERIC Document Reproduction Service, #ED 054 233.)

Goodman, K. S. "Behind the Eye: What Happens in Reading." In *Reading: Process and Program*, edited by K. S. Goodman and O. S. Niles. Urbana, Illinois: National Council of Teachers of English, 1970, 1-38.

Harris, A. J. *Effective Teaching of Reading*. New York: David McKay, 1962.

Johnson, N. F. "Influence of Associations Between Elements of Structural Verbal Responses." *Journal of Verbal Learning and Verbal Behavior* 5 (1966): 369-74.

Lorge, I. "Readability Formulae—An Evaluation." *Elementary English* 26 (1949): 229–33.

MacGinitie, W. H. *Contextual Constraint in English Prose*. Ph.D. dissertation, Teachers College, Columbia University, 1960.

Miller, G. A. "Some Psychological Studies of Grammar." *American Psychologist* 17 (1962): 748–62.

Potter, T. C. *A Taxonomy of Cloze Research, Part I: Readability and Reading Comprehension*. Southwest Regional Laboratory for Education Research and Development, Technical Report No. 11. Los Angeles, 1968.

Rankin, E. J. "Cloze Procedure—A Survey of Research." *Yearbook of the Southwest Regional Conference* 14 (1965): 133–48.

5 Diagnostic Procedures
Auditory Discrimination: Differences versus Deficits

Peggy E. Williams

The recent past witnessed an increased interest in the learning processes of the so-called disadvantaged child. This interest resulted from the discovery that vast discrepancies existed between the academic achievement of what the literature calls "lower class" children, especially "lower class" black children, and the achievement of middle and upper class youngsters. To investigate this lag, researchers descended upon the communities of the "deprived." Their goal was to explain the lag and to construct improved models for educating "deprived" youngsters. Their task was to describe and to prescribe.

The descriptions resulting from these *investigations* became enormously important in the subsequent development of educational programs for "lower class" children. When a black child was described as "culturally deprived," there was an assumption that he or she had no language, no customs, no social institutions, no behavior patterns, and no expectations for the future. The resulting educational program sought to fill this cultural void through inculcation of the language, the customs, and the behavior patterns of the dominant culture. On the other hand, if a black youngster is described as "culturally different," the assumption is that he or she has a language, customs, social institutions, behavior patterns, and expectations which are shared with other members of the cultural group; a very different kind of educational program evolves from the need to prepare youngsters to function in a society which acknowledges and respects the contributions of all of its citizens.

If there are to be viable programs for educating the less fortunate members of this society, the importance of accurate descriptions of their learning difficulties cannot be underestimated. Differences must be clearly distinguished from deficits.

The development and standardization of the *Wepman Auditory Discrimination Test* or *WADT* (Wepman 1958), along with Wep-

man's findings which indicated a significant relationship between poor articulation and auditory discrimination and between auditory discrimination and reading achievement (1960), stimulated a great deal of interest in the auditory discrimination capacity of "lower class" black youngsters who had been described by educators as inarticulate non-readers.

The *WADT* is an individually administered test which consists of forty word pairs. Thirteen word pairs differ in initial consonant phoneme; thirteen differ in final consonant phoneme; four differ in medial vowel phoneme; and ten are the same. The youngster's task is to listen as the word pairs are read by the examiner and to indicate whether the pairs are the "same" or "different" (Wepman 1958).

The present inquiry was inspired by the author's experiences in the diagnosis and remediation of reading problems and by the reports of deficit theorists who argue that the environment of "the lower class" inhibits the development of auditory discrimination, thereby creating in "lower class" children deficiencies in auditory discrimination ability. Of particular interest was the early research of Cynthia Deutsch (1967a) which apparently set the stage for subsequent studies relating auditory discrimination ability to social class and race.

Deutsch reported that there was a significant correlation between reading ability and auditory discrimination in black "lower class" males. (These findings supported Wepman's data.) Deutsch (1967a, pp. 268-69) also reported a significant relationship between verbal performance and performance on the *WADT*.

Deutsch alluded to animal studies which indicated that auditory-evoked potential is reduced when the animal is attentive to stimuli in other modalities. These findings, according to Deutsch, are consistent with findings that signal-to-noise ratio influences the stimulus perceived and the response evoked. She concluded that accurate perception results from a higher signal-to-noise ratio. Deutsch hypothesized that with a low ratio, or a lot of noise in the system, excessive activation of the reticular system interferes with the travel of the signal up neural paths. She further hypothesized that excessive activation beyond the optimal level necessary to prepare the organism to accept and respond to a stimulus produced a blocking out of the stimulus. Since auditory stimuli are ever present, Deutsch postulated that auditory stimuli are particularly prone to a "tuning-out" process or learned inattention. Deutsch (1968, p. 68) reported that the "tuning-out" process did not take place at the level of the reticular system; rather the

stimuli reaches the central projection area but is not attended to. She contended that one could expect that a child raised in a very noisy environment with little directed and sustained speech stimulation would be deficient in his or her discrimination and recognition of speech sounds. One could further expect, according to Deutsch, that the child raised in a noisy environment would be relatively inattentive to auditory stimuli and have difficulty with other skills that depend upon good auditory discrimination. She suggests that the slum child lives in a very noisy environment and gets little connected or concentrated speech directed to him or her. Deutsch (1967a, pp. 262–75) concluded that if her hypotheses are accepted and viewed in light of her research findings (significant correlations between verbal ability and auditory discrimination and reading ability and auditory discrimination), it could well be that "lower class" youngsters, specifically "lower class" black youngsters, who live in noisy environments do not develop the requisite auditory discrimination ability to learn to read well. Middle class children from quieter and more speech-directed environments do not have this problem.

Deutsch's arguments are reported here in detail because it would appear that they engendered the idea that auditory discrimination deficiencies might be characteristic of certain ethnic groups and socioeconomic levels. Her hypotheses were widely accepted, oft-quoted, and served as a basis for much of the subsequent research on auditory discrimination. Clark and Richards (1966, p. 262) reported that their research revealed pre-school economically disadvantaged children exhibited significant deficiencies in auditory discrimination. Berlin and Dill (1967, p. 386) tested the effects of feedback and reinforcement on "lower class" black children's performance on the *WADT*. Oakland (1969, p. 37) found that a significant relationship between auditory discrimination and reading existed when reading achievement scores of "lower class" children were correlated with the results of a phonemic test of auditory discrimination (*WADT*), but the relationship did not exist when reading achievement was correlated with the results of a non-phonemic assessment of auditory discrimination.

Deficit theorists reported a significant difference between the auditory discrimination ability of "lower" and "middle class" children. The results of their studies revealed that poor readers within social class groups experienced significantly more difficulty in auditory discrimination than good readers and that the difference between good and poor readers was greater for the "lower class" group. In support of Cynthia Deutsch's tentative hypotheses,

Martin Deutsch (1967a, p. 48) suggested the possibility that the difficulties encountered by "lower class" children could be attributed to noisy homes, lack of practice in auditory discrimination, and limited verbal interaction between children and adults.

Bereiter and Engelmann (1966, pp. 30–35) compared "lower class" black children to deaf children and implied that auditory problems caused them to treat strings of single words as though they were one word. They called this "the giant-word syndrome."

Although Oakland (1969) sought to add to the support for a correlation between reading achievement and auditory discrimination among "disadvantaged" youngsters, he recognized the possibility that ability to perform well on a test of auditory discrimination might be affected by the dialect of the examiner. Linguists contend that the dialect of the youngster might also affect performance on a phonemic test of auditory discrimination. Labov (1969b, p. 57) observed in his studies of the language of black youth gangs that black youngsters experienced difficulty perceiving phonemic contrasts which were not in their own dialect. Baratz (1969b, p. 47) has indicated that whereas the test most widely used to measure auditory discrimination (the *WADT*) equates correct responses with judgments of equivalences and differences in "standard" English sound usage, black children respond on the basis of the sound usage learned in their social and geographical milieu. According to Johnson (1971, p. 151), phonological conflict points exist between the "standard" sound usage and the sound usage of black youngsters. Johnson concluded that because these youngsters are accustomed to a different sound system, they are therefore equipped with different auditory skills.

The arguments advanced by Labov, Baratz, and Johnson are supported by the research of Gross (1967, p. 2124), who investigated the relationship between reading comprehension and dialect divergence and between pronunciation and auditory discrimination. Gross first investigated the sounds omitted or "distorted" by his sample of 72 black youngsters. His findings revealed that the sounds most frequently omitted or changed were also the sounds most frequently missed in auditory discrimination. Gross's investigation further revealed that the subjects experienced more difficulty discriminating word pairs comprised of words which had an alternate dialect pronunciation (example: wind/win').

Stevens (1972, p. 66) has suggested that tools for collecting information about various ethnic and cultural groups are not functionally equivalent across groups. If research data are to be

meaningful, investigatory procedures must control for critical differences between various groups. The deficit theorists relied almost exclusively on the *WADT* to measure auditory discrimination ability. The validity of the deficit theory is contingent upon the appropriateness of the *WADT* as an instrument for measuring the auditory discrimination ability of youngsters whose dialects differ from the dialect upon which the test is based. Validity also depends upon the research procedures used to control for the dialect differences. Speakers of the "nonstandard" dialects might perform poorly on the *WADT* because several items reflect phonological conflict points (Politzer 1971, p. 177). This investigation was undertaken to study several of these items and to determine their effect on performance.

Procedure

Subjects. The subjects for this study consisted of 48 black youngsters who were nearing the end of fifth grade. To assure variance in the speech patterns represented in the sample, subjects were drawn from two very diverse populations. Prerequisites for participation included evidence of the absence of hearing defects and ability to read and comprehend the passage used to assess speech patterns. Only average and above average readers were chosen. The sample consisted of 22 boys and 26 girls ranging in age from 10 to 12 years.

Measuring Instruments. Two informal instruments were constructed to identify speech patterns. A reading passage was developed which was loaded with words that lent themselves to *e/i, f/th* and *v/the* neutralizations which, according to the literature, are characteristic of the speech of the "lower class." These neutralizations are reflected in items on the *WADT—pen/pin* (item 40), *sheaf/sheath* (item 28) and *clothe/clove* (item 25) (Wepman 1958). The passage was followed by seven comprehension questions which were read to the subject by the examiner. The ten responses elicited required the subject to repeat selected words. The second instrument used a sound-matching task to further assess the presence of neutralizations. The instruments which evaluated speech patterns consisted of 65 items or words which reflected the aforementioned neutralizations.

To establish the reliability of the informal measures of speech patterns, these 65 items were compared to a Rasch Model to determine those which best measured speech patterns. The Rasch Model is a special form of the logistic response latent traits model and is used to select those items in a test which are good for

measuring a specific trait (Wright 1967). This analysis indicated the 15 items which best measured speech patterns. These items showed a reliability of +.81 (N=48). A speech pattern score was tabulated based on the 15 items and this score was used in subsequent analyses.

Form I of the *WADT* (Wepman 1958) was used to measure development in auditory discrimination.

Results. Pearson product-moment correlations were computed between *WADT* errors and errors on items 25, 28, and 40 of the *WADT*. These items correspond with *v/the*, *f/th*, and *e/i* neutralizations and the speech pattern score previously discussed. The results indicated no significant relationship between *e/i* neutralizations and item 40 of the *WADT*. Since all of the subjects neutralized *pen* and *pin*, failure to establish a positive correlation here is doubtless due to the lack of variance among the subjects with respect to this particular neutralization pattern. Subsequent analyses did not include examination of the relationship between item 40 and *e/i* neutralizations.

Supporting the hypothesis that youngsters would experience difficulty discriminating items which reflect phonological conflict points, the correlation between *WADT* errors and errors on items reflecting conflict points (*clothe/clove* and *sheaf/sheath*) was significant at the $p<.01$ level.

In order to ascertain whether the *WADT* score and performance on items 25 and 28 (*clothe/clove* and *sheaf/sheath*) measured speech patterns as well as auditory discrimination, a multivariate analysis of variance was done. The speech pattern scores were divided into low, average, and high groups. Of interest here was whether the *WADT* scores of the three groups were significantly different. The differences among the three groups' performance on the *WADT* and on items 25 and 28 (*clothe/clove* and *sheaf/sheath*) was significant at the $p<.01$ level, supporting the hypothesis that speech patterns are predictive of performance on the *WADT*

A subsequent analysis of co-variance was done to determine the influence that performance on items 25 and 28 had on the total *WADT* score. The results of the univariate analysis of co-variance indicated that items 25 and 28 were significantly related to total *WADT* score ($p<.01$); however, there was no significant difference ($p<.23$) between the performance of the low, average, and high speech pattern groups when those items were removed from the test

Discussion

Deficit theorists used comparative analyses of the test scores and language behavior of "lower" and "middle class" youngsters to conclude that "lower class" black children were linguistically and auditorily deficient. Linguists, in the main, dealt with descriptions of syntactic and phonological systems and avoided hierarchical judgments about the effectiveness of one system over another. They concluded that while the syntactic and phonological systems of "lower class" black youngsters may differ in some respects from "standard" English, these differences do not constitute deficiencies.

There is little empirical evidence to support either position on auditory discrimination deficits versus differences. Evidence used in defense of the deficit theory appears to be based on an auditory discrimination task which apparently measures more than ability to discriminate pairs of phonemes.

Cynthia Deutsch (1967a) postulated the existence of auditory discrimination deficiencies resulting from an environment which inhibits development of auditory discrimination. Deutsch's hypotheses are based on a significant correlation between reading achievement scores and performance on the *WADT*. There was no data presented correlating the actual living conditions of the subjects with performance on the *WADT* in the statistical analyses offered to support the hypotheses.

The present study documents the difficulties that can result from *WADT* items which reflect conflict points between the dialect of the subject and the dialect of the test. This data supports the findings of linguists, as well as the observations of deficit theorists who admit that "lower class" youngsters perform as well as "middle class" youngsters in discriminating initial phonemes (although their ability to distinguish final phonemes is significantly different).

The task of judging the sameness or difference of word pairs is complex. It involves hearing, perceiving, memorizing, and decision making (Blank 1968, p. 1100; Politzer 1971). Assuming that all children have an equal chance of overcoming most of the complexities inherent in this kind of test, one factor continues to separate "middle class" youngsters, who very likely speak the dialect of the test, from those who do not, specifically black, Spanish-speaking, and Appalachian white youngsters. That separating factor is the difference between the phonological systems of these children and the system upon which the test, the *WADT* in this case, is based.

The factors which influence performance on the *WADT* must be

controlled if it is to be validly used in research. Deficit theorists did not control for dialect differences; therefore, deficit hypotheses are invalidated by their dependence on the *WADT*.

It is important to note that Wepman (1960 and 1968) defined the appropriate use of the *WADT*. It is apparent from his definition that the test was constructed to measure individual development in order to more effectively meet individual needs. When used as prescribed, where responses can be evaluated in terms of influencing factors, the *WADT* is a good diagnostic tool. The use of *WADT* scores to generalize group deficiencies and to justify group training to cure group ills is a misapplication of the test. Wepman warned that group auditory training is ineffectual.

Effective educational programs can evolve only from precise research data. Historically, programs resulting from inaccurate descriptions have consumed considerable time and money but failed to extensively improve the lot of the "underprivileged" child. Labov (1969a, p. 72) contends that the interventionalists, in their evaluation of the data, will not fault the theory or the method but will fault the child. Those who framed the Headstart Programs were suggesting that at age four it was already too late to reverse the effects of cumulative deficit; therefore, "deprived" youngsters must be reached even sooner.

Cynthia Deutsch (1967b), despite Wepman's warnings, attempted group auditory training with a group of third grade reading retardates who were deficient in the "skill" of auditory discrimination. After a semester of training, there was no differential auditory discrimination improvement between those who received specific training and those who did not.

Because the deficit theorists assumed that performance on the *WADT* was functionally equivalent across ethnic groups, they failed to control for differences between the phonological systems of their subjects and that upon which the *WADT* was structured. Their failure to use controls appears to invalidate their findings. Future research should establish different stimulus situations for different ethnic groups so that comparable behavior can be observed. Reports based on the supposition of a white middle class norm are meaningless for the optimal development of youngsters who do not fit this category.

REFERENCES

Baratz, Joan C., and Roger W. Shuy, editors. *Teaching Black Children to Read.* Washington, D.C.: Center for Applied Linguistics, 1969.

Bereiter, Carl, and Siegfried Engelmann. *Teaching Disadvantaged Children in the Preschool.* Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1966.

Berlin, Charles I., and Anne C. Dill. "The Effects of Feedback and Positive Reinforcement on the Wepman Auditory Discrimination Test Scores of Lower-Class Negro and White Children." *Journal of Speech and Hearing Research* 10 (June 1967): 384-89.

Blank, Marion. "Cognitive Processes in Auditory Discrimination in Normal and Retarded Readers." *Child Development* 39 (December 1968): 1091-1101.

Clark, Ann D., and Charlotte J. Richards. "Auditory Discrimination among Economically Disadvantaged and Nondisadvantaged Pre-school Children." *Exceptional Children* 33 (December 1966): 259-62.

Deutsch, Cynthia P. "Auditory Discrimination and Learning: Social Factors." In *The Disadvantaged Child,* edited by Martin Deutsch and Associates. New York: Basic Books, Inc., 1967a, 259-77.

——. "Learning in the Disadvantaged." In *The Disadvantaged Child,* edited by Martin Deutsch and Associates. New York: Basic Books, Inc., 1967b. 147-62.

——. "Environment and Perception." In *Social Class, Race, and Psychological Development,* edited by Martin Deutsch, et al. New York: Holt, Rinehart and Winston, Inc., 1968, 56-85.

Deutsch, Martin. "The Disadvantaged Child and the Learning Process." In *The Disadvantaged Child,* edited by Martin Deutsch and Associates. New York: Basic Books, Inc., 1967a, 39-58.

——. "Social Intervention and the Malleability of the Child." In *The Disadvantaged Child,* edited by Martin Deutsch and Associates. New York: Basic Books, Inc., 1967b. 3-30.

Gross, Reuben S. "Dialect Pronunciation, Auditory Discrimination and Reading." Ph.D. dissertation. *Dissertation Abstracts* 28B (October-December 1967): 2124.

Johnson, Kenneth R. "The Influence of Nonstandard Negro Dialect on Reading Achievement." *The English Record* XXI/4 (April 1971): 148-54.

Labov, William. "The Logic of Non-Standard English." *The Florida FL Reporter* 7 (Spring/Summer 1969a): 60-74.

——. "Some Sources of Reading Problems for Negro Speakers of Nonstandard English." *Teaching Black Children to Read,* edited by Joan C. Baratz and Roger W. Shuy. Washington, D.C.: Center for Applied Linguistics, 1969b, 29-66.

Oakland, Thomas D. "Auditory Discrimination and Socioeconomic Status as Correlates of Reading Ability." *Journal of Learning Disabilities* 2 (June 1969): 33-37

Politzer, Robert L. "Auditory Discrimination and the 'Disadvantaged': Deficit or Difference." *The English Record* XXI, 4 (April 1971): 174-78.

Stevens, Joseph H., Jr. "Current Directions in the Study of Parental Facilitation of Children's Cognitive Development." *Educational Horizons* 50, 2 (Winter 1972): 62-67.

Wepman, Joseph M. *Manual of Directions: Auditory Discrimination Test*. Chicago: Joseph M. Wepman, 1958.

———. "Auditory Discrimination, Speech and Reading." *Elementary School Journal* 60/6 (March 1960): 325-33.

———. "Auditory Discrimination: Its Role in Language Comprehension, Formulation and Use." *Pediatric Clinics of North America* 15/3 (August 1968): 721-77.

Wright, Benjamin D. *Sample-free Test Calibration and Person Measurement*. New Jersey: Educational Testing Service, 1967.

Use of Informal Reading Inventories

William D. Page and Rebecca C. Barr

Teaching, as well as clinical work in reading, involves insightful observation and analysis of the child's performance in order to plan, implement, and evaluate effective learning situations. The informal reading inventory may be the most useful observation technique available for assessing reading performance in order to plan effective instruction for readers who have not done well in conventional situations (Farr 1969, p. 98). For those unfamiliar with the informal reading inventory, Johnson and Kress (1965) provide a concise, practical introduction. Some ideas from research warrant application when informal reading inventories are used; in this review, we will describe selected insights and knowledge from research and draw implications for using informal reading inventories.

Teachers who are able to observe, analyze, and interpret oral reading responses possess one of the most useful skills for assessing children's reading. From understanding gained about students' reading processes, a teacher can plan appropriate instruction and evaluate its effectiveness. In recent years, the considerable research using oral reading analysis as a basic tool has furnished new insights. Suggestions for improving informal reading analysis fall into four areas: (1) using and selecting passages, (2) recording oral reading responses, (3) interpreting oral reading responses, and (4) evaluating comprehension.

Using and Selecting Passages

Sources of Graded Paragraphs. It is common knowledge that a number of commercially prepared publications provide a variety of graded paragraphs bound for long-term use (Harris 1970, pp. 138–47). By drawing upon several published inventories, teachers can provide themselves with several sets of graded paragraphs.

Published inventories may also provide norms or average expected scores at specific grade levels.

Teacher-constructed inventories composed of selections from basal readers or workbooks have the added advantage of enabling the teacher to identify appropriate instructional materials in addition to determining what reading strategies the child uses. A number of researchers have described the use of graded textbooks for assessing reading performance by having children try to read various books. Workbook passages may be preferable to basal passages when they include prepared questions for evaluating comprehension. Passages can be graded with the level assigned to books or workbooks by the publishers, through traditional readability methods, or more quickly and reliably through cloze readability procedures described by Bormuth (1974).

A clinician or teacher who is trained in statistics, or has statistical consultant help, can generate norms or average score expectations for passages specifically for the children who make up the instructional population. The point is that concerned teachers or clinicians need not be bound to less productive procedures because of a lack of graded paragraphs.

Using Graded Paragraphs. Graded paragraphs from standardized individual reading inventories can be readily used in at least three ways. One use involves a listening test where the clinician reads a passage to the subject and asks comprehension questions; this yields a listening comprehension score which can index how well a child may be able to learn to read.

A second use of graded paragraphs involves an oral reading test. This test yields (1) an oral reading word recognition score, (2) miscues—unanticipated oral responses (K. Goodman 1965)—that can be analyzed to reveal how the reader uses contextual and word recognition strategies in comprehending print, (3) an oral reading comprehension score, and (4) a reading rate score. The third use involves a silent paragraph reading test, which yields a reading rate score and a silent reading comprehension score. All of these tests provide a clinician or teacher opportunity to observe the child while reading. To fully use the test tactics—listening comprehension, oral reading, and silent reading—three sets of graded paragraphs are required. Presently, few commercial inventories supply more than two sets of graded paragraphs. Clinicians need not abandon their favorite inventory, but should simply augment it with additional sets of graded paragraphs.

Longer Passages. Menosky (1971) demonstrated that readers process differently at the beginning of a passage than they do at the middle or end. Changes occurred with her sample around the 250- or 300-word mark. It was not clear whether the changes were due to fatigue, learning resulting from processing, or other factors. It is clear that we can learn much about a reader's processing in oral reading by using passages that substantially exceed 300 words. It is also important to note the correspondence between normal reading demands in school and test passage length. Frequently, in normal classroom activities, students are required to read passages that exceed the length of graded paragraphs used in most commercially prepared inventories. The widespread practice of using paragraphs of about 200 words or less in informal reading inventories need not be abandoned; instead, we can augment present inventories with a longer passage.

To obtain a sufficient number of responses for analysis, the longer passage must be reasonably difficult. Performance on shorter paragraphs can provide the information needed to select a longer passage that is appropriate. The longer passage should be about one level more difficult than the instructional level established by conventional informal reading inventory procedures (see Johnson and Kress 1965 for conventional procedures). By selecting a set of longer passages—500 to 600 words—that correspond to the levels in the informal inventory, a relatively efficient technique can be added to present informal reading inventories.

Recording Oral Reading Responses

Observing Oral Reading. Experiences of miscue researchers caused reexamination of the observer's role in both reading research and clinical work. Three language processing interactions are sources of variation in observing oral reading (Page 1973, p. 68). One interaction involves the reader processing the print. The two other language processing interactions involve the clinician or teacher, who (a) also processes the print and produces an internal reaction to it that functions as an expected response and (b) processes the reader's oral responses and produces an internal reaction that functions as the observed response. The three interactions—the reader to print, the clinician to print, and the clinician to the reader's responses—are interdependent. Each interaction involves the use of language and each is a source of variation in observing oral reading performance.

Dialect. Present practice generally ignores the two interactions involving the clinician or teacher. Though researchers have not yet taken up the challenge that these interactions pose, teachers and reading clinicians can apply the insights in several ways. When a clinician's dialect is distinctly different from a reader's dialect, it is reasonable for the clinician or teacher to carefully examine expectations and the reader's performance. The reader's responses may be acceptable in his or her dialect even if they deviate from the clinician's dialect. Experienced clinicians have long recognized this problem and many have intuitively adjusted their assessment techniques. Evidence suggests that dialect divergence probably does not interfere with comprehension (Goodman and Buck 1973; Sims 1972). When weighing oral reading error scores against comprehension scores for a dialect divergent reader, the comprehension score should take precedence, since differences in dialect between the clinician and the reader may have inflated the oral reading errors.

Tape Recording. Another implication of the three interactions involves memory. Both of the clinician's internal responses, the response to the print and the response to the reader's oral response, must be held in memory while the clinician compares them. The widespread practice of marking a typed script of the passage while the reader responds should be continued, but an additional inexpensive technique, already commonly used, can reduce the heavy reliance on the clinician's memory. By tape recording the reader's oral responses, the clinician can analyze the responses as many times as necessary. The clinician can also get help from others in analyzing the responses, a procedure that researchers use regularly to establish observer reliability. This is particularly helpful when the dialects of the clinician and reader diverge extremely. Miscue research has capitalized on tape recording and demonstrated that it yields substantially more information than can be gained by listening once as the reader reads (K. Goodman and Burke 1973).

Interpreting Oral Reading Responses

Language Knowledge. A third implication of the oral reading observation process described above involves recognizing the part that language plays when anyone analyzes oral reading performance. A speaker of English can do far more than simply identify and count oral reading errors, since the observer can use his or her language to assess the type and quality of the response. Miscue research provides an empirical framework and the *Reading Miscue*

Inventory (Y. Goodman and Burke 1972) provides an organizational structure for applying the insights in clinical work. The language of the author, the reader, and the clinician or teacher are inseparably involved. Though linguistics is purely descriptive and cannot tell us how language should be, the descriptions of how language works that linguists provide can improve our understanding of what happens in the reading process.

K. Goodman (1965) studied oral reading errors as miscues or deviations from the text rather than as mistakes indicating poor reading. He constructed a model that depicted reading as a sampling process in which all the available information is not used, but only selected cues are processed to form guesses (1970, pp. 30-31). Goodman (1967) described reading as an inquiry process in which the reader seeks to verify or reject his or her guesses on the basis of information and inferences drawn from three cue systems: the semantic, syntactic, and graphophonic. The semantic cue system draws on both word meaning and contextual meaning. The syntactic cue system involves grammatical structures and the arrangement of words, phrases and clauses, etc. The graphophonic cue system relates the sounds and graphic forms of printed language (Allen 1972).

Substitution and Mispronunciation Miscues. Insights can be gained by analyzing miscues that are substitutions or mispronunciations. It is useful to separate substitution responses into two groups: (1) frequently used content and function words found in basic lists, and (2) unusual words. Two questions help us analyze the substitutions or mispronunciations produced in response to frequently used words. First, what proportion of the total number of sight words in the passage generated miscues? Second, do these miscues distort the author's meaning? If very few of the total substitution and pronunciation miscues are responses to frequently used words, and those miscues do not distort the author's meaning, then they are probably no cause for concern. However, if the portion of substitutions or mispronunciations with frequently used sight words is high, or if the miscues distort the author's meaning, further diagnostic work and teaching of sight words may be indicated. Isolated sight word recognition should not be construed as a test of the full reading process; rather it indicates development in one aspect of reading. Similarly, instruction in sight word development should not be limited to isolated words. Experience stories and activities using words in a variety of contexts provide the most direct and logically defensible route to reading with understanding.

How a child identifies words can be inferred by analyzing the child's substitutions and mispronunciations in response to unusual content words. The method used to teach the child to read will influence how he or she attacks unknown words (Elder 1971; DeLawter 1970). Typically, a child who has been taught reading with primary emphasis on comprehension and the development of a sight vocabulary produces meaningful substitutions and few mispronunciations that are nonwords. The substitution miscues tend to approximate the printed word in length; initial consonant(s) and sometimes final consonant(s) usually agree with the text. The majority of substitutions are semantically appropriate, at least within the immediate sentence context. There may be a tendency to correct semantically inappropriate words.

In contrast, the oral reading responses of a child taught with emphasis on phonic decoding tend to display the following miscue characteristics: a high proportion of mispronunciations or nonwords, a high degree of graphophonic similarity between the printed word and the oral response, a tendency for a high proportion of the miscues not to be semantically appropriate, and correction made mainly on the basis of graphic rather than contextual cues. The nature of a child's functional reading strategies has implications for instruction. If a child's listening comprehension exceeds reading comprehension by a grade or more, his or her present reading strategies can probably be improved through instruction.

The Correction Process. The number and type of miscues that prompt the reader to make corrections illuminate how he or she is processing the print. Some miscues elicit *no* correction attempts because the reader is satisfied. Some defy correction because the reader has no means for recognizing the particular miscue within his or her system of processing print; this may reflect instructional practice. Still other miscues appear to go uncorrected because the reader chooses not to reveal recognition of the mistake.

A reader who corrects a response is evidencing dissatisfaction with the original oral response. These decisions to correct can be sorted into several categories. Consider an oral reading miscue that is semantically and syntactically acceptable but differs graphophonically from the expected response, as in "Tim ran down the *road*" for "Tim ran down the *lane*." If the reader does not attempt to correct the miscue, it is reasonable to assume that the semantic and syntactic cue systems permitted him or her to verify the guess and feel satisfied with the original response. We can also assume that, to some degree, the graphophonic discrepancy did not cause

dissatisfaction. If no pauses or other miscues immediately precede or follow this response, the assumption about a weak graphophonic system is reinforced. If the reader uses *road* and *lane* interchangeably in a semantically acceptable way throughout the passage, this assumption is strengthened further still.

If, on the other hand, the reader attempts to correct the response of *road* for *lane*, we may assume the graphophonic cue systems are operating because the original response caused dissatisfaction. An attempt to correct may be successful or unsuccessful. If successful, and no further correction attempts occur, we may assume the graphophonic cue system to be operating well. If the successful correction is abandoned through a subsequent unsuccessful attempt, we have less support for assuming the graphophonic cue system is operating.

A semantically and syntactically unacceptable response, such as "Tim ran down the *lone*," warrants interpretation in a similar vein. No correction attempt suggests that the semantic, syntactic, and graphophonic cue systems are all weak to some degree. If the reader attempts to correct and struggles with a vowel change such as *lean* for *lane*, we can assume the graphophonic system is generating the correction attempt. On the other hand, a correction attempt of *line* for *lane* suggests all cue systems may be involved.

When a reader abandons an originally correct response, as some do, and attempts one or several corrections, we can theorize about the cue systems that are operating. Clay (1967 and 1968) found low error but high self-correction with high-ability 5-year-olds, while low-ability subjects produced a high incidence of errors and low self-correction. K. Goodman and Burke (1973, p. 90), with a sample ranging from 2nd to 10th grade, found low-ability readers corrected graphically uncompliant miscues more consistently than did average- and high-ability readers, who also tended not to correct phonemically and graphemically highly compliant miscues. Clearly, a careful look at which miscues elicit a correction and which can go uncorrected can yield a great deal of information about the reading process, information that present practices tend to ignore.

Evaluating Comprehension

Difficulties with measuring reading comprehension are not limited to informal reading inventories. Problems center about inconsistencies in attempts to measure assumed subskills of reading comprehension (Farr 1969, p. 56) and inadequate attempts to define comprehension entirely in terms of visible activities. As

Simons says, ". . . the comprehension process is inaccessible to direct observation . . . (1971, p. 340)." In view of these difficulties, one should not use results from a single test of comprehension without corroborative information. "A test should be viewed as only one piece of information for making decisions (Farr and Roser 1974, p. 598)." Additional information on comprehension performance can extend the web of interrelated information on a child's reading performance and thus provide a better base for educational prescription; descriptions follow of three tactics that can be used to augment conventional comprehension tests.

Retelling Comprehension. Y. Goodman and Burke (1972) have explained how teachers can assess comprehension by having the subject retell the story and respond to a series of open-ended questions. The procedures can be applied to passages 300 words and longer. Analysis of the retelling generates a comprehension rating. K. Goodman and Burke (1973) provided empirical backing for the relationship of the retelling rating to oral reading miscues.

The Comprehending Score. A single passage of 500 to 600 words permits assessing comprehension with a "comprehending score," using procedures outlined by Rousch (1972) and K. Goodman and Burke (1973). A comprehending score was defined as: the percent of oral reading miscues that are semantically acceptable added to the percent of miscues that are semantically unacceptable but successfully corrected. The comprehending score correlated highly with the retelling comprehension rating in the study by K. Goodman and Burke (1973, p. 68). No criterion norms are yet available for the comprehending score.

Comprehending scores are simple and inexpensive to calculate. As is the case with most scores, they do not yield dependable information alone, but used in conjunction with other scores, they help fill in a picture of the reader's processing. Since the comprehending score is generated entirely from oral reading performance and the retelling comprehension rating is generated from the child's responses following oral reading, the two scores corroborate each other. In addition, a comprehending score can be computed on the first half of a 500- to 600-word passage and compared with the comprehending score on the second half to reveal changes in processing after 250 words (Menosky 1971).

The Cloze Test. A cloze test with every fifth word deleted and exact answers scored correct has high concurrent validity with multiple choice tests of literal comprehension (Bormuth 1969, p. 365). This validity and the fact that it is reliable, inexpensive,

and easily administered make the cloze test a very reasonable technique to incorporate in the informal reading inventory. The cloze test needs 50 or more items; therefore, passages of 250 words or more are required. Bormuth (1972 and in this volume) provides criteria for interpreting cloze test performance. The cloze test warrants a great deal of further study because it is convenient and reliable and it provides another tool to help determine what is actually measured by conventional comprehension tests. Clearly, we need to develop new tools to build more adequate explanations of comprehension than presently exist.

Conclusion

The ideas in this paper begin to show us the improvements in informal reading inventories that we can make with insights from recent research. Researchers have much to do in the areas of validity and measurement. The comprehending score warrants further study. The interpretation of miscues and their corrections promises better ways of assessing performance. Publishers may be expected to provide more graded paragraphs, both short and long. More precise indices of conventional cloze performance can be expected. The concerned reading teacher, reading clinician, and administrator have much to gain by following the breakthroughs that regularly occur in current research, with the ultimate goal of better helping youngsters learn to read.

REFERENCES

Allen, Paul David. "Cue Systems Available during the Reading Process: A Psycholinguistic Viewpoint." *Elementary School Journal* 72/5 (February 1972): 258-64.

Bormuth, John. "Factor Validity of Cloze Tests as Measures of Reading Comprehension Ability." *Reading Research Quarterly* IV/3 (Spring 1969): 358-65.

——. "Literacy in the Classroom." Paper presented at the University of Chicago Graduate School of Education Summer Reading Conference 1972. (Included in this volume.)

Clay, Marie. "The Reading Behavior of Five-Year-Old Children: A Research Report." *New Zealand Journal of Educational Studies* 2 (May 1967): 11-31.

——. "A Syntactic Analysis of Reading Errors." *Journal of Verbal Learning and Behavior* 7 (1968): 434-38.

DeLawter, Jayne. "Oral Reading Errors of Second Grade Children Exposed to Two Different Reading Approaches." Ph.D. dissertation, Teachers College, Columbia University, 1970.

Elder, Richard. "Oral Reading Achievement of Scottish and American Children." *Elementary School Journal* 71:4 (January 1971): 216-30.

Farr, Roger. *Reading: What Can Be Measured.* Newark, Delaware: International Reading Association, 1969.

Farr, Roger and Nancy Roser. "Reading Assessment: A Look at Problems and Issues." *Journal of Reading* 17:8 (May 1974): 592-99.

Harris, Albert. *How to Increase Reading Ability.* New York: David McKay Inc., 1970.

Goodman, Kenneth S. "A Linguistic Study of Cues and Miscues in Reading." *Elementary English* 42 (1965): 639-43.

———. "Reading: A Psycholinguistic Guessing Game." *Journal of the Reading Specialist* 6 (1967): 126-35.

———. "Behind the Eye: What Happens in Reading." In *Reading: Process and Program*, edited by Kenneth S. Goodman and Olive S. Niles. Urbana, Illinois: National Council of Teachers of English, 1970, 1-38.

Goodman, Kenneth S., and Catherine Buck. "Dialect Barriers to Reading Comprehension Revisited." *The Reading Teacher* 27 (1973): 6-12.

Goodman, Kenneth S., and Carolyn L. Burke. *Theoretically Based Studies of Patterns of Miscues in Oral Reading Performance.* USOE Project No. OEG0-9-320375-4269. Washington, D.C.: U.S. Department of Health, Education, and Welfare, March 1973.

Goodman, Yetta, and Carolyn L. Burke. *Reading Miscue Inventory.* Toronto, Canada: Macmillan Publishing Company, Inc.

Johnson, Mariorie, and Roy Kress. *Informal Reading Inventories.* Newark, Delaware: International Reading Association, 1965.

Menosky, Dorothy. "A Psycholinguistic Description of Oral Reading Miscues Generated during the Reading of Varying Portions of Text by Selected Readers from Grades Two, Four, Six, and Eight." Ph.D. dissertation, Wayne State University, 1971.

Page, William D. "Clinical Uses of Miscue Research." In *Miscue Analysis: Applications to Reading Instruction*, edited by Kenneth S. Goodman. Urbana, Illinois: National Council of Teachers of English and ERIC Clearinghouse on Reading and Communication Skills, 1973, 65-76.

Rousch, Peter D. "A Psycholinguistic Investigation into Oral and Silent Reading, Prior Conceptual Knowledge and Post-Reading Performance: A Study in Miscue Analysis." Ph.D. dissertation, Wayne State University, 1972.

Simons, Herbert D. "Reading Comprehension: The Need for a New Perspective." *Reading Research Quarterly* 6 (Spring 1971): 338-63.

Sims, Rudine. "A Psycholinguistic Description of Miscues Generated by Selected Young Readers during the Oral Reading of Text Materials in Black Dialect and Standard English." Ph.D. dissertation, Wayne State University, 1972.